



**Cátia Isabel  
Assis Fidalgo**

**Comunidades bacterianas endofíticas de *Halimione  
portulacoides***

**Endophytic bacterial communities of *Halimione  
portulacoides***





**Cátia Isabel  
Assis Fidalgo**

**Comunidades bacterianas endofíticas de  
*Halimione portulacoides***

**Endophytic bacterial communities of *Halimione  
portulacoides***

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Biologia, realizada sob a orientação científica do Doutor Artur Jorge da Costa Peixoto Alves, Investigador Principal do Departamento de Biologia da Universidade de Aveiro, e da Doutora Isabel da Silva Henriques, Investigadora Auxiliar do Departamento de Biologia da Universidade de Aveiro

Apoio financeiro da FCT e do FEDER  
através do programa COMPETE no âmbito  
do projeto de investigação PhytoMarsh.

Bolsas com referência:  
PTDC/AAC-AMB/118873/2010  
FCOMP-01-0124-FEDER-019328

Apoio financeiro da Fundação para a  
Ciência e Tecnologia e do Fundo Social  
Europeu no âmbito do III Quadro  
Comunitário de Apoio.

Bolsa de Doutoramento:  
SFRH/BD/85423/2012



*It's not easy being green*



## **o júri**

presidente

Doutor Armando da Costa Duarte  
Professor Catedrático da Universidade de Aveiro

vogais

Doutora Paula Maria Lima e Castro  
Professora Catedrática da Universidade Católica Portuguesa

Doutora Martha Estela Trujillo Toledo  
Professora Titular da Universidade de Salamanca, Espanha

Doutora Paula Maria de Melim e Vasconcelos de Vitorino Morais  
Professora Auxiliar da Universidade de Coimbra

Doutora Maria Ângela Sousa Dias Alves Cunha  
Professora Auxiliar da Universidade de Aveiro

Doutor Artur Jorge da Costa Peixoto Alves  
Investigador Principal da Universidade de Aveiro





## **agradecimentos**

Começo por agradecer ao meu orientador Doutor Artur Alves e co-orientadora Doutora Isabel Henriques pelos conhecimentos que me transmitiram, e pelo que me ajudaram ao longo destes últimos anos. Nada disto seria possível sem o vosso empenho, interesse e paciência.

Aos co-autores dos trabalhos desenvolvidos, agradeço pelos vossos importantes contributos: Jaqueline Rocha, Ricardo Martins, Raúl Riesco, Doutora Marta Tacão, Doutor Diogo Proença, Professora Martha Trujillo e Professora Paula Morais.

Aos restantes colaboradores do projecto PhytoMarsh, Doutora Alexandra Moura, Doutora Cláudia Oliveira, Professora Etelvina Figueira e Professor António Correia, agradeço pelos variados contributos durante o projecto. Ao último, Professor António Correia, reforço ainda um agradecimento por me ter acolhido no grupo microlab.

A todo o microlab, agradeço por me terem acolhido no grupo, pelo espírito de equipa e entreajuda, e pela permanente troca de conhecimentos que me ajudou a crescer.

Aos colegas do microlab e fora deste, especialmente Carina, Marta Alves, Susana, Liliana, Zé, Anabela, Patrícia, Laura, Forough, Eliana, Carla, Helena Correia, Helena Lopes, Fernanda, Rafael Félix, Paulo, Cláudio, Rafael Tavares, Sofia Libório, Isabel Silva, Joana Amaral e Bárbara Correia, o maior dos agradecimentos pela ajuda, amizade e companheirismo.

Aos outros colegas do departamento, agradeço pelas vossas contribuições para este trabalho: Cátia Velez, Doutor Bruno Castro.

Aos funcionários do Departamento de Biologia e CESAM, especialmente a Dona Helena, Engenheiro Armando, Engenheiro Fernando Cozinheiro e Doutora Anabela Pereira, agradeço a disponibilidade e a ajuda nas mais variadas situações.

À minha família, em especial aos meus pais e irmão, pelo apoio que sempre me deram e por sempre terem acreditado em mim.

Finalmente, à Zara e à Maria.



## palavras-chave

Bactérias endofíticas, *Halimione portulacoides*, promoção do crescimento de plantas, diversidade bacteriana, espécies bacterianas novas, sequenciação de próxima geração, Illumina.

## resumo

Os sapais são ecossistemas marinhos altamente produtivos que frequentemente recebem contaminantes de natureza antropogénica. A Ria de Aveiro encontra-se no noroeste de Portugal e contém numerosos sapais. *Halimione portulacoides* é um dos halófitos mais importantes em sapais Europeus e tem sido amplamente estudada devido ao seu potencial para ser usada em fins de fitorremediação, e como bioindicador de contaminação de sedimentos. Bactérias endofíticas podem apresentar capacidade promotora do crescimento de plantas (PCP), quer diretamente por produção de fito-hormonas e aquisição de nutrientes, quer indiretamente via competição com fitopatogenos. No presente trabalho, a diversidade de bactérias endofíticas da planta de sapal *H. portulacoides* da Ria de Aveiro é explorada extensivamente.

Isolados de bactérias endofíticas foram obtidos e caracterizados quanto à sua taxonomia, capacidade de produzir enzimas e características PCP. As características mais observadas foram atividade celulolítica, xilanolítica e desaminase de 1-aminociclopropano-1-carboxilato, e a produção da auxina ácido indol-3-acético. Os resultados revelaram um enorme potencial da coleção para PCP *in vitro* e *in vivo*.

A coleção de isolados foi também explorada para procurar diversidade não descrita. Como resultado, dez novas espécies de bactérias foram amplamente caracterizadas e descritas: *Microbacterium diaminobutyricum*, *Saccharospirillum correiae*, *Altererythrobacter halimionae*, *Altererythrobacter endophyticus*, *Zunongwangia endophytica*, *Salinicola halimionae*, *Salinicola aestuarina*, *Salinicola endophytica*, *Salinicola halophytica* e *Salinicola lusitana*. Consequentemente, o presente trabalho expôs a endosfera de *H. portulacoides* como um foco de diversidade bacteriana desconhecida.

A composição taxonómica da comunidade endofítica foi averiguada via sequenciação do gene 16S rRNA da coleção de isolados, e mais profundamente com a utilização de sequenciação de alto rendimento independente do cultivo. A última abordagem revelou cinco filos principais: Proteobacteria, Planctomycetes, Actinobacteria, Bacteroidetes e Firmicutes. Destes, apenas Planctomycetes não foi obtido na coleção de isolados. As comunidades diferiram de acordo com o local (no ensaio dependente do cultivo, para locais contaminados e não-contaminado) e tecido (em ambos os ensaios) de amostragem. As principais famílias obtidas no endofitoma nuclear foram *Oceanospirillaceae* em tecidos de parte aérea, e *Enterobacteriaceae* e *Kiloniellaceae* em tecidos de raiz.

O trabalho apresentado providenciou uma compreensão profunda das bactérias endofíticas presentes no halófito *H. portulacoides*, e expôs o seu potencial como foco de bactérias não descritas e bactérias promotoras do crescimento de plantas.



## keywords

Endophytic bacteria, *Halimione portulacoides*, plant growth promotion, bacterial diversity, novel bacterial species, next-generation sequencing, Illumina.

## abstract

Salt marshes are highly productive marine ecosystems that often act as a sink for contaminants of anthropogenic nature. The Ria de Aveiro lagoon is located in the north-west of Portugal and comprises numerous salt marshes. *Halimione portulacoides* is one of the most important halophytes in European salt marshes and has been widely researched for its potential for phytoremediation, and as a bioindicator of sediment contamination. Endophytic bacteria can present plant growth promotion (PGP) abilities, either directly by production of phytohormones and nutrient uptake, or indirectly via competition with phytopathogens. In the present work, the diversity of endophytic bacteria from the salt marsh plant *H. portulacoides* from Ria de Aveiro is extensively explored.

Endophytic bacterial isolates were obtained and characterized for their taxonomy, ability to produce specific enzymes and PGP traits. The most observed traits were cellulolytic, xylanolytic and 1-aminocyclopropane-1-carboxylate deaminase activities, and the production of the auxin indol-3-acetic acid. The results revealed an enormous potential of the collection for *in vitro* and *in vivo* PGP.

The collection of isolates was also explored for undescribed diversity. As a result, ten novel bacterial species were thoroughly characterized and described: *Microbacterium diaminobutyricum*, *Saccharospirillum correae*, *Altererythrobacter halimionae*, *Altererythrobacter endophyticus*, *Zunongwangia endophytica*, *Salinicola halimionae*, *Salinicola aestuarina*, *Salinicola endophytica*, *Salinicola halophytica* and *Salinicola lusitana*. Consequently, the present work exposes the endosphere of *H. portulacoides* as a hotspot of unknown bacterial diversity.

The taxonomic composition of the endophytic community was assessed via 16S rRNA gene sequencing of the isolate collection, and with more depth using culture-independent high-throughput sequencing. The latter approach revealed five main phyla: Proteobacteria, Planctomycetes, Actinobacteria, Bacteroidetes and Firmicutes. From these, only Planctomycetes was not obtained in the isolate collection. The communities differed according to sampling site (for the culture-dependent assay, for contaminated and non-contaminated sites) and tissue (in both assays). The main families found in the core endophytome were *Oceanospirillaceae* for aboveground tissues, and *Enterobacteriaceae* and *Kiloniellaceae* for belowground tissues.

The present work provided a deep understanding of the endophytic bacteria present in the halophyte *H. portulacoides*, and exposed its potential as a hotspot of undescribed bacteria and plant growth promoting bacteria.



## Table of contents

List of figures.....	v
List of tables.....	vii
<b>Chapter 1: Introduction .....</b>	<b>9</b>
1.1 Plant species: <i>Halimione portulacoides</i> .....	11
1.2 Study area: Ria de Aveiro and Laranjo Bay .....	13
1.3 Endophytic bacteria .....	16
1.4 References .....	23
<b>Chapter 2: Objectives.....</b>	<b>27</b>
2.1 Scope .....	29
2.2 Main goal .....	30
2.3 Specific goals .....	30
2.4 Hypotheses .....	30
<b>Chapter 3: Culturable endophytic bacteria from <i>Halimione portulacoides</i>.....</b>	<b>31</b>
3.1 Culturable endophytic bacteria from the salt marsh plant <i>Halimione portulacoides</i> : phylogenetic diversity, functional characterization, and influence of metal(loid) contamination .....	33
Abstract .....	34
Background .....	35
Methods.....	37
Results .....	44
Discussion .....	52
Acknowledgments .....	58
References .....	59
Supplementary Material.....	65
3.2 Diversity of endophytic <i>Pseudomonas</i> in <i>Halimione portulacoides</i> from metal(loid)-polluted salt marshes.....	84
Abstract .....	85
Background .....	86
Methods.....	88
Results and Discussion .....	94
Conclusions .....	105
Acknowledgements .....	105
References .....	106
Supplementary Material .....	112

<b>Chapter 4: Novel bacterial species.....</b>	<b>115</b>
<b>4.1 <i>Microbacterium diaminobutyricum</i> sp. nov., isolated from <i>Halimione portulacoides</i>, which contains diaminobutyric acid in its cell wall, and emended description of the genus <i>Microbacterium</i> .....</b>	<b>117</b>
Abstract .....	118
Main text.....	119
Emended description of the genus <i>Microbacterium</i> Orla-Jensen 1919 Takeuchi and Hatano 1998 ...	129
Description of <i>Microbacterium diaminobutyricum</i> sp. nov. ....	129
Acknowledgements .....	130
References .....	131
Supplementary material .....	135
<b>4.2 <i>Saccharospirillum correiae</i> sp. nov., an endophytic bacterium isolated from the halophyte <i>Halimione portulacoides</i>.....</b>	<b>138</b>
Abstract .....	139
Main text.....	140
Description of <i>Saccharospirillum correiae</i> sp. nov. ....	146
Acknowledgements .....	148
References .....	148
Supplementary material .....	150
<b>4.3 <i>Altererythrobacter halimionae</i> sp. nov. and <i>Altererythrobacter endophyticus</i> sp. nov., two endophytes from the salt marsh plant <i>Halimione portulacoides</i> .....</b>	<b>151</b>
Abstract .....	152
Main text.....	153
Description of <i>Altererythrobacter halimionae</i> sp. nov. ....	159
Description of <i>Altererythrobacter endophyticus</i> sp. nov. ....	160
Acknowledgements .....	161
References .....	162
Supplementary material .....	164
<b>4.4 <i>Zunongwangia endophytica</i> sp. nov., an endophyte isolated from the salt marsh plant <i>Halimione portulacoides</i>, and emended description of the genus <i>Zunongwangia</i>.....</b>	<b>166</b>
Abstract .....	167
Main text.....	168
Emended description of the genus <i>Zunongwangia</i> Qin et al. 2007, emend. Rameshkumar et al. 2014 .....	174
Description of <i>Zunongwangia endophytica</i> sp. nov. ....	174
Acknowledgements .....	175
References .....	176



<b>4.5 The endosphere of the salt marsh plant <i>Halimione portulacoides</i> is a diversity hotspot for the genus <i>Salinicola</i>: description of five novel species <i>Salinicola halimionae</i> sp. nov., <i>Salinicola aestuarina</i> sp. nov., <i>Salinicola endophytica</i> sp. nov., <i>Salinicola halophytica</i> sp. nov. and <i>Salinicola lusitana</i> sp. nov. ....</b>	<b>178</b>
Abstract .....	179
Background .....	180
Motivation .....	183
Methods.....	183
Results and discussion .....	186
Emended description of the genus <i>Salinicola</i> Anan'ina et al. 2007, emend. de la Haba et al. 2010b..	203
Description of <i>Salinicola halimionae</i> sp. nov. CPA60 <sup>T</sup> .....	203
Description of <i>Salinicola aestuarina</i> sp. nov. CPA62 <sup>T</sup> .....	204
Description of <i>Salinicola endophytica</i> sp. nov. CPA92 <sup>T</sup> .....	205
Description of <i>Salinicola halophytica</i> sp. nov. CR45 <sup>T</sup> .....	206
Description of <i>Salinicola lusitana</i> sp. nov. CR50 <sup>T</sup> .....	207
Acknowledgements .....	209
References .....	209
<b>Chapter 5: Illumina-based analysis of endophytic bacteria from <i>Halimione portulacoides</i> .....</b>	<b>215</b>
Background .....	217
Motivation .....	219
Methods.....	220
Results and discussion .....	225
Conclusions .....	260
References .....	261
Annex I .....	267
Annex II .....	268
<b>Chapter 6: Final considerations .....</b>	<b>281</b>
Final considerations .....	283
Conclusions .....	289
Future work.....	289
References .....	290



## List of figures

<b>Figure 1.1</b> Schematics of the phytosphere.....	16
<b>Figure 3.1</b> Distribution of 658 endophytic isolates by class amongst isolation sites (B, E and C) and plant tissues .....	47
<b>Figure 3.2</b> Principal coordinates analysis (PCoA) plot based on 16S rRNA gene partial sequences from twenty one endophytic bacterial samples from all isolation sites and tissues .....	48
<b>Figure 3.3</b> Distribution of isolates that tested positive for enzymatic activity assays and plant growth promotion traits from 467 representative isolates.....	51
<b>Figure 3.4</b> Neighbor-joining (bootstrap=1000) tree illustrating the phylogenetic position of the <i>Pseudomonas</i> isolates obtained in this study and the closest related <i>Pseudomonas</i> type strains .....	96
<b>Figure 3.5</b> Cluster analysis of the antibiotic susceptibility profiles of <i>Pseudomonas</i> isolates from <i>H. portulacoides</i> , using Bray-Curtis similarity coefficient and unweighted pair group method, using arithmetic averages cluster methods .....	98
<b>Figure 3.6</b> Cluster analysis of plasmidic DNA restriction profiles from endophytic <i>Pseudomonas</i> isolates..	100
<b>Figure 3.7</b> Effects of endophytic <i>Pseudomonas</i> inoculation on <i>A. thaliana</i> seedlings growth .....	104
<b>Supplementary Figure S3.1</b> Distribution of isolates that tested positive for enzymatic activity assays and plant growth promotion traits from 467 representative isolates .....	83
<b>Figure 4.1</b> UPGMA cluster analysis based on the Pearson coefficient of ERIC-PCR fingerprints of <i>Microbacterium</i> spp. strains isolated from roots of <i>Halimione portulacoides</i> .....	122
<b>Figure 4.2</b> Maximum-likelihood phylogenetic tree based on concatenated sequences of <i>gyrB</i> , <i>rpoB</i> , <i>recA</i> and <i>ppk</i> and 16S rRNA genes, showing the relationship between strains isolated from <i>Halimione portulacoides</i> and a set of type strains of species of the genus <i>Microbacterium</i> .....	123
<b>Figure 4.3</b> Neighbour-joining (NJ) tree showing the phylogenetic positions of strain CPA1 <sup>T</sup> and representatives of other related taxa based on 16S rRNA gene sequences .....	141
<b>Figure 4.4</b> Neighbour-joining (NJ) tree showing the phylogenetic positions of strains CPA5 <sup>T</sup> and BR75 <sup>T</sup> and representatives of other related taxa based on 16S rRNA gene sequences .....	155
<b>Figure 4.5</b> Neighbour-joining (NJ) tree showing the phylogenetic positions of strain CPA58 <sup>T</sup> and representatives of other related taxa based on 16S rRNA gene sequences .....	170
<b>Figure 4.6</b> Neighbour-joining (NJ) phylogenetic trees showing the phylogenetic positions of strains CPA60 <sup>T</sup> , CPA62 <sup>T</sup> , CPA92 <sup>T</sup> , RZ23, CR45 <sup>T</sup> , CR50 <sup>T</sup> , CR57 and representatives of other related taxa based on 16S rRNA gene (a), <i>gyrB</i> (b) and <i>rpoD</i> (c) sequences .....	195
<b>Figure 4.7</b> Neighbor-joining (NJ) phylogenetic tree showing the phylogenetic positions of strains CPA60 <sup>T</sup> , CPA62 <sup>T</sup> , CPA92 <sup>T</sup> , RZ23, CR45 <sup>T</sup> , CR50 <sup>T</sup> , CR57 and representatives of other related taxa based on the concatenated sequences of the 16S rRNA, <i>gyrB</i> and <i>rpoD</i> genes .....	196

<b>Supplementary Figure S4.1</b> Maximum-likelihood phylogenetic tree based on nearly full-length 16S rRNA gene sequences, showing the relationship between strains isolated from <i>Halimione portulacoides</i> and type strains of the genus <i>Microbacterium</i> .....	135
<b>Supplementary Figure S4.2</b> Polar lipid profile of strain RZ63 <sup>T</sup> after staining with: A, ninhydrin, B, molybdotophosphoric acid and C, anisaldehyde.....	136
<b>Supplementary Figure S4.3</b> Two-dimensional thin-layer chromatography of the total polar lipids of strain CPA1 <sup>T</sup> (A) and <i>Saccharospirillum impatiens</i> DSM 12546 <sup>T</sup> (B) stained with 5% ethanolic molybdophosphoric acid .....	150
<b>Supplementary Figure S4.4</b> Neighbour-joining (NJ) phylogenetic tree based on nearly-full length 16S rRNA gene sequences of strains CPA5 <sup>T</sup> and BR75 <sup>T</sup> , isolated from <i>Halimione portulacoides</i> , and type strains of the genus <i>Altererythrobacter</i> , other genera of the family <i>Erythrobacteraceae</i> , and relevant species of the <i>Novosphingobium</i> genus .....	165
<b>Supplementary Figure S4.5</b> Two-dimensional TLC of total polar lipids of strain CPA5 <sup>T</sup> (A), BR75 <sup>T</sup> (B), <i>A. aestuarii</i> KCTC 22735T (C) and <i>A. marensis</i> KCTC 22370T (D) stained with 5% ethanolic molybdophosphoric acid .....	165
 <b>Figure 5.1</b> Schematics of the sampling strategy to obtain fifteen specimens of <i>Halimione portulacoides</i> from Ria de Aveiro .....	220
<b>Figure 5.2</b> Rarefaction curves for all samples .....	233
<b>Figure 5.3</b> Alpha-diversity metrics results.....	234
<b>Figure 5.4</b> Boxplot of total number of reads per OTU for the original and transformed OTU tables .....	236
<b>Figure 5.5</b> Non-metric multidimensional scaling (NMDS) of the endophytic community according to plot and tissue of origin .....	239
<b>Figure 5.6</b> Principal coordinates analysis (PCoA) of the endophytic community according to plot and tissue of origin.....	240
<b>Figure 5.7</b> Venn diagrams built using OTUs from the endophytic bacterial community .....	242
<b>Figure 5.8</b> Taxonomic visualization of the endophytic bacterial community from <i>H. portulacoides</i> .....	245
<b>Figure 5.9</b> Read abundance for the twenty top families according to sampling tissue and plot.....	250
<b>Figure 5.10</b> Read abundance for the top core families for each tissue.....	253

## List of tables

<b>Table 3.1</b> Metal(loid) concentrations in sediment and plant tissue samples.....	44
<b>Table 3.2</b> Analysis of 16S rRNA gene partial sequences with RDP pipeline .....	48
<b>Table 3.3</b> Screening of bacterial endophytes for enzymatic activity and plant growth promotion traits .....	49
<b>Table 3.4</b> Plant growth promotion traits, antimicrobial, and enzymatic activities of <i>Pseudomonas</i> isolates from <i>H. portulacoides</i> .....	102
 <b>Supplementary Table S3.1</b> Colony forming units (CFUs) per gram of vegetal tissue (fresh weight) in samples from all isolation sites (C, E and B) and tissues, per culture media .....	65
<b>Supplementary Table S3.2</b> Phylum, class and genus-level distribution of 658 bacterial endophytic isolates in three sites (C, E and B) and two types of tissues.....	67
<b>Supplementary Table S3.3</b> Taxonomic attribution, accession number and results for enzymatic activity and plant growth promotion assays of 467 representative isolates .....	70
<b>Supplementary Table S3.4</b> PCR primers and conditions for antibiotic and metal resistance genes amplification .....	112
 <b>Table 4.1</b> Primer sets developed and PCR amplification conditions for each gene .....	121
<b>Table 4.2</b> Differential characteristics of strains obtained from <i>Halimione portulacoides</i> and phylogenetically related <i>Microbacterium</i> species, <i>M. flavum</i> DSM 18909 <sup>T</sup> and <i>M. lacus</i> DSM 18910 <sup>T</sup> .....	125
<b>Table 4.3</b> Chemotaxonomic markers of selected members of the genus <i>Microbacterium</i> .....	128
<b>Table 4.4</b> Cellular fatty acid profiles of strains isolated from roots of <i>Halimione portulacoides</i> and phylogenetically related species of the genus <i>Microbacterium</i> .....	129
<b>Table 4.5</b> Differential characteristics of strain CPA1 <sup>T</sup> and species of the genus <i>Saccharospirillum</i> with validly published names .....	143
<b>Table 4.6</b> Fatty acid composition of CPA1 <sup>T</sup> and type strains of species of the genus <i>Saccharospirillum</i> .....	145
<b>Table 4.7</b> Differential characteristics of strains CPA5 <sup>T</sup> , BR75 <sup>T</sup> , and related type strains (data from this study unless otherwise stated) .....	157
<b>Table 4.8</b> Fatty acid composition of strains CPA5 <sup>T</sup> , BR75 <sup>T</sup> and related type strains. ....	158
<b>Table 4.9</b> Physiological characteristics of strain CPA58 <sup>T</sup> and recognized species of the genus <i>Zunongwangia</i> (data from this study unless otherwise stated).....	172
<b>Table 4.10</b> Fatty acid composition of strains CPA58 <sup>T</sup> , <i>Z. profunda</i> DSM 18752 <sup>T</sup> , <i>Z. mangrovi</i> DSM 24499 <sup>T</sup> . 173	
<b>Table 4.11</b> Information of sequenced genomes .....	187
<b>Table 4.12</b> Length of 16S rRNA gene sequences obtained from RAST annotation and best matches observed by comparison against the EzTaxon database (Kim et al., 2012) .....	189
<b>Table 4.13</b> Pairwise comparison of genome sequence-based information of the endophytic strains and related type strains of the genus <i>Salinicola</i> .....	191
<b>Table 4.14</b> Length and origin of sequences used in the MLSA approach.....	193
<b>Table 4.15</b> Complete fatty acid composition of the endophytic strains CPA60 <sup>T</sup> , CPA62 <sup>T</sup> , CPA92 <sup>T</sup> , CR45 <sup>T</sup> and CR50 <sup>T</sup> , and related type strains .....	198

<b>Table 4.16</b> Differential and phenotypic characteristics of strains CPA60 <sup>T</sup> , CPA62 <sup>T</sup> , CPA92 <sup>T</sup> , CR45 <sup>T</sup> and CR50 <sup>T</sup> , and related type strains.....	200
<b>Supplementary Table S4.1</b> GenBank accession numbers of <i>gyrB</i> , <i>rpoB</i> , <i>recA</i> , and <i>ppk</i> genes.....	136
<b>Table 5.1</b> Description of the samples used in the study .....	221
<b>Table 5.2</b> Number of reads after each sequence data processing step .....	226
<b>Table 5.3</b> Comparison of number of reads resulting from PCR amplification in the absence and presence of PNA blockers .....	230
<b>Table 5.4</b> Alpha-diversity metrics results .....	232
<b>Table 5.5</b> Coverage estimation values reported for Illumina-based studies of endophytic bacterial communities.....	232
<b>Table 5.6</b> Comparison of average values of alpha-diversity metrics from culture-dependent and culture-independent methodologies .....	235
<b>Table 5.7</b> Phyla of plant-associated bacteria observed in different halophyte species.....	246
<b>Table 5.8</b> Families of plant-associated bacteria observed in closely related halophyte species .....	248

# Chapter 1

---

Introduction

## **Contents**

**1.1** Plant species: *Halimione portulacoides*

**1.2** Study area: Ria de Aveiro and Laranjo Bay

**1.3** Endophytic bacteria



### 1.1 Plant species: *Halimione portulacoides*

*Halimione portulacoides* (L.) Aellen (basonym *Atriplex portulacoides*, synonym *Obione portulacoides*; The International Plant Names Index, 2012), common name sea purslane, is an eudicot from the order Caryophyllales, family *Amaranthaceae*, sub-family *Chenopodioideae*. This C3 evergreen small shrub can reach 80 cm in height, can be found in acid, neutral and basic soils, and can grow with little or no shade conditions (Anjum et al., 2014). Leaves from this plant can be used for human consumption, either uncooked in salads, or cooked as a potherb, and its use for asthma treatment has been reported (Everest & Ozturk, 2005). With a world wide distribution, *H. portulacoides* is the most abundant, and one of the most productive species in ungrazed European salt marshes (Bouchard et al., 1998; Anjum et al., 2014).

Halophytes are often considered for salt marsh phytoremediation of metals and other contaminants, since they represent a relatively inexpensive option for storing contaminants in their biomass (Figueira et al., 2012). In addition to *H. portulacoides*, other halophytes such as *Spartina alterniflora*, *Juncus maritimus*, and *Phragmites australis* have been considered for their phytoremediation potential (Williams et al., 1994; Figueira et al., 2012; Amari et al., 2017). The use of *H. portulacoides* in phytoremediation of contaminated sediments has been largely researched in several Portuguese marshes, such as Ria de Aveiro, Tagus, Sado, and Guadiana marshes (Canário et al., 2007).

The potential for phytoremediation of *H. portulacoides* has been particularly explored for mercury (Hg) contaminated sediments. In fact, the ability of this halophyte to be a biomonitor for Hg contamination has been explored, since this halophyte is able to give quantitative information regarding Hg contamination levels in the sediments (Válega et al., 2008a). Hg is accumulated in belowground (BG) tissues of the halophyte, and its concentration in aboveground (AG) tissues is comparatively low. Some reasons have been given for this difference in Hg content in both tissues: (i) Hg mobility occurs primarily between the sediment and BG tissues, and there is a poor translocation of this metal from BG to AG tissues (Canário et al., 2007; Válega et al., 2008b; Caçador et al., 2009; Canário et al., 2017), (ii) there is a constant renovation of leaves thus keeping metal content low at any given moment (Válega et al., 2008a), (iii) there is low accumulation in stems since they are transport and not storage organs (Válega et al., 2008a), (iv) leaves of the halophyte are able to exude Hg<sup>0</sup> from the AG tissues to the atmosphere (Canário et al., 2017). Moreover, retaining metals in the BG tissues may be a strategy to protect the more sensitive AG tissues from

the effects of the toxic metals (Lozano-Rodriguez et al., 1997). The accumulation of Hg in the BG tissues causes the halophyte to act as a sink for the toxic element, however, Hg translocation from BG to AG tissues, albeit low, may act as a potential Hg source in salt marshes (Anjum et al., 2011).

The discrepancy in metal accumulation in the different plant tissues has also been observed for other metals: copper (Cu), zinc (Zn), cadmium (Cd; Caçador et al., 2009). For cobalt (Co), however, it was observed that up to 50 % of the total Co content was contained in AG tissues (Caçador et al., 2009). Seasonal variation for Cu, Zn, Cd and Co accumulation was observed in AG tissues, but not in BG tissues (Caçador et al., 2009). Seasonality was also verified for Hg accumulation in *H. portulacoides* (Anjum et al., 2011).

Metal tolerance in *H. portulacoides* may be associated with the accumulation of these toxic elements in the cell walls in both BG and AG tissues, preventing the contaminants to exert their effects in the cytoplasm (Válega et al., 2009). Sequestration of Hg into vacuoles via cysteine-rich peptides is another mechanism for metal tolerance exhibited by salt marsh plants (Canário et al., 2007 and references therein). Recently, phytochelatins have been suggested to have a role in tolerance to arsenic (As), Zn and lead (Pb) in salt marsh plants under natural conditions (Negrin et al., 2017).

The ability for *H. portulacoides* to remediate sediments contaminated with other types of contaminants has also been examined, namely for hydrocarbons (Almeida et al., 2008; Couto et al., 2011) and tributyltin (Carvalho et al., 2010). Phytoremediation of Cu by salt marsh plants including *H. portulacoides* is affected by the presence of polycyclic aromatic hydrocarbons (Almeida et al., 2008) and by whether the metal is in ionic or nanoparticle (NP) form (Andreotti et al., 2015; Fernandes et al., 2016). The NP form may result in aggregates which lead to lower Cu accumulation, thus reducing metal availability and uptake. Consequently, the phytoremediation potential is diminished (Fernandes et al., 2016). These additional contaminants and form of contaminant should be taken into account when analyzing the potential for phytoremediation in contaminated sediments.

## **1.2 Study area: Ria de Aveiro and Laranjo Bay**

Salt marshes are highly productive estuarine ecosystems characterized by high salinity and periodical flooding, providing conditions for specialized colonization. In addition to these restrictive conditions, salt marshes are often affected by anthropogenic activities and are considered a sink for contaminants, including metals (reviewed in Williams et al., 1994).

The Ria de Aveiro lagoon, located in the northwest coast of Portugal alongside the Atlantic Ocean, is 45 km long and 10 km wide, and contains salt marshes that are subjected to semi-diurnal tides (Válega et al., 2008b). This shallow lagoon contains several channels and extensive inter-tidal areas and is connected to the Atlantic Ocean through a narrow opening known as Barra de Aveiro (Lopes et al., 2001; Anjum et al., 2014).

For over four decades (1950-1994), an industrial complex in Estarreja discharged contaminant-heavy effluents to the Ria de Aveiro lagoon, through a remote branch of the lagoon that ends at the Laranjo Bay. In 1994, the reduction of contaminated discharges was achieved with changes on production processes, installation of a recovery plant, a wastewater treatment plant, and a landfill for solid waste and contaminated sediments (Costa and Jesus-Rydin, 2001; Pereira et al., 1998b). Main contaminants present in the effluents discharged to Laranjo Bay included aniline, benzene, monochlorobenzene, mononitrobenzene, As, Hg, Zn and Pb. Those present in higher concentrations were As, Hg, Pb and Zn, where As and Hg were the most severe and spread (Costa and Jesus-Rydin, 2001). Notwithstanding the diversity of contaminants present in Laranjo Bay, particular attention has been given to Hg and its contents in different elements of the ecosystem. No biological function is known for Hg, and it is considered a priority by the European Water Framework Directive ([www.http://ec.europa.eu/environment/water/water-framework](http://ec.europa.eu/environment/water/water-framework)) and on a global scale, since events of bioconcentration, bioaccumulation and biomagnification can be observed when this contaminant is present in waters (Pereira et al., 2009 and references therein).

The Laranjo Bay is a lagoon connected to the Ria de Aveiro through a narrow channel. The sediment at Laranjo Bay is a mixture of sand and mud containing 35-75 % of fine particles (Válega et al., 2008b), and exhibits a marked north-south distribution of cohesive sediments. The northern area contains more cohesive sediments which allow for an easier adsorption of metallic elements, resulting in the formation of stable metal-sediment complexes (Lopes et al., 2001).

The Hg content of the water column is low (10 to 37 pmol dm<sup>-3</sup>) in comparison with pore waters in the sediments (17 to 188 pmol dm<sup>-3</sup>). The stability of the complexes formed between the metal and the sediments prevents the enrichment of the water column (Ramalhosa et al., 2006). Analysis of the vertical profiles of Laranjo Bay sediments revealed the lowest Hg content at the core bottom, an increase and peak of Hg contents in layers at 10 to 25 cm, and a decrease toward the surface sediments. These profiles reflect a pre-industrial period, followed by the contaminant-heavy discharges, and the more recent reduction of intake of industrial contaminants (Pereira et al., 1998a; Pereira et al., 1998b; Valega et al., 2008b). Analysis of Hg content in the water column and in the sediments show that the highest concentrations are observed closest to the source of the contamination, and a contamination gradient is observed inside the lagoon (Pereira et al., 1998b; Ramalhosa et al., 2006; Valega et al., 2009).

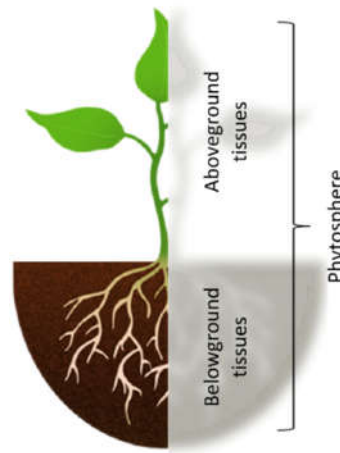
In 1998 it was estimated that a 33 ton load of Hg was present in the Ria de Aveiro lagoon, and that most of this load was confined to the Estarreja Channel and Laranjo Bay: approximately 27 ton in less than 4 % of the Ria de Aveiro area (Pereira et al., 1998a; Pereira et al., 1998b). There is evidence that Hg is continuously being exported from the Laranjo Bay, especially through tidal forces by resuspension of the contaminated sediments in the water column (Pereira et al., 1998b; Lopes et al., 2001). The surface layers of the sediments are more easily eroded and contaminants transferred to the water column present a potential risk for spread of the contamination throughout the network of the Ria de Aveiro (Lopes et al., 2001; Pato et al., 2008b). Deeper layers of sediments, which are richer in contaminant content, are not as easily eroded to the water column, unless events such as floods and storms are observed. These high energy phenomena are able to resuspend Hg-rich particles existent in the sediments and to transfer pore waters to the water column (Pato et al., 2008a). This erosion of contaminated sediments and consequent export does not, however, contribute to a major contamination problem for nearshore sites, since the semi-enclosed structure of the Laranjo Bay prevents these Hg-associated particles from being transported outside the Bay (Lopes et al., 2001). In addition, the complex network of channels in the Ria de Aveiro lagoon captures contaminated particules before coastal waters are reached; and nearby marine sediments have a low ability to incorporate Hg, reducing the impact on the coastal zone (Pato et al., 2008b).

Vegetated sediments are distinct from non-vegetated sediments in the Laranjo Bay: the vegetated sediment is more acidic, presents higher redox potential and higher organic matter content (Válega et al., 2008c), and is more contaminated with Hg than non-vegetated sediment (Válega et al., 2008b). The latter is due to the ability of salt marsh plants to act as sediment traps, contributing to an elevated Hg content in the sediments, even after the reduction in contaminated effluents from the industrial complex (Válega et al., 2008c).

### 1.3 Endophytic bacteria

#### The phytosphere

The term phytosphere refers to the entirety of the tissues of a plant, and can be divided into different compartments (Figure 1.1) such as (i) aboveground (AG) tissues or phyllosphere (comprising stems, leaves, flowers); and (ii) belowground (BG) tissues or rhizosphere (root system). For both compartments, two divisions can be considered: the endosphere (inner tissues) and the episphere or ectosphere (outer tissues).



**Figure 1.1** Schematics of the phytosphere.

All plants in natural settings have, at least in a phase of their life cycle, association with microorganisms (Partida-Martínez & Heil, 2011). Bacteria, archaea, fungi, algae and amoeba have been found in the different compartments of the phytosphere, revealing a microbial hotspot in this complex system (reviewed in Hardoim et al., 2015). Vandenkoornhuyse et al. (2015) discussed in detail the perspective of the plant as a holobiont, as a dynamic entity composed by the plant host and its microbiome, working together toward evolutionary fitness. The authors expose that the holobiont as a whole is the target of evolutionary processes, and that, consequently, the holobiont as a unit adapts to environmental settings. As a result of this adaption, plant-associated microbes are involved in functions such as improving the nutrition of the host and the resistance to biotic and abiotic stresses (Vandenkoornhuyse et al., 2015). Plant-associated bacteria are known to possess these functions, as they can participate in metal tolerance (Canário et al., 2007) and degradation of contaminants (Carvalho et al., 2010; Oliveira et al., 2014a; Oliveira et al., 2014b), among many other plant-beneficial role.

## **Endophytic bacteria**

### Defining endophytic bacteria

The definition of endophyte has changed throughout the years. In 1926, Perotti described endophyte as a particular stage in the life of a bacteria, as an advanced stage of infection, which is similar to mutualistic symbiosis (Perotti, 1926). Endophytes were then considered all those microorganisms that could be isolated from surface-sterilized plant organs (Henning & Villforth, 1940). In 1992, Quispel introduced the concept of endophytic bacteria as those that colonize the inner tissues of the plant without damaging the host or eliciting strong defense responses. The latter definition of endophytic bacteria is currently widely used and is the definition that will be adopted throughout the present work.

Bacterial endophytes may present different relationships with their plant hosts in the sense that there can be opportunistic or facultative endophytes. The former relates to endophytes that spend part of their life cycle as an epiphyte and occasionally migrate to inner tissues. The second refers to those that exhibit a more mutualistic life style with the plant host. An additional type of endophytic relationship is the 'obligate' endophyte, where there is a requirement of plant tissues to complete the microorganism's life cycle. However, obligate endophytes have only been reported for fungi (Hardoim et al., 2015 and references therein).

Regarding the function of the endophytic bacteria in the plant host, three groups can be considered: commensal, plant growth promoters or latent pathogens. The first group lives on the metabolites produced by the plant but have no apparent effect on the host. The second group causes a positive impact on the host by protecting against phytopathogens and exhibiting plant growth promoting (PGP) traits, as detailed below (Hardoim et al., 2015). Consequently, the microbiota of a plant can be seen as facilitators in the sense that the host-associated microbes provide additional genes that can be useful to adapt to environmental conditions (reviewed in Vandenkoornhuyse et al., 2015).

Two types of endophytes may be considered according to their host-specificity. The first type comprises bacteria that bind to plant root surfaces in a non-specific manner and then migrate to inner tissues via wounds, cracks and lenticels in the root system. The second type comprises bacteria that specifically bind to certain plant species, e.g., rhizobia that bind to particular plant species to form nodules (Glick, 2015). Endophytes of the first type are, nevertheless, subjected to some degree of specificity due to the fact that different plant species produce different exudates that attract certain bacteria (Hardoim et al., 2015).

### Colonization of inner tissues

Bacteria can enroll in the endophytic lifestyle either through horizontal or vertical transmission.

Horizontal transmission results from the entry of bacteria from the external environment into the plant, and this type of colonization most often occurs as bacteria enter the root system from the rhizosphere. The most common way of entry occurs via primary and lateral root hair cells, root cracks, wounds and hydrolysis of root cells. This hydrolysis process may be related to the production of enzymes with cellulolytic activity, however, colonization still occurs in the absence of cellulases (reviewed in Santoyo et al., 2016). Entry can also take place at the root tips through germinating radicles and openings that result from the natural growth process of the roots, or through lenticels (Hardoim et al., 2008; Santoyo et al., 2016). Colonization by specific groups of bacteria in the root system can result in the formation of root nodules, where nitrogen fixation can occur. Additional entry points also arise in AG tissues of plants, namely through stomata, lenticels and wounds caused by other organisms such as insects and fungi, hydathodes, flowers and fruits (reviewed in Hardoim et al., 2015; Santoyo et al., 2016). Bacterial taxa present in the ectosphere and in the endosphere may differ significantly, suggesting that the endosphere colonizers are equipped with a set of genes that allow for their establishment in the inner host tissues (reviewed in Reinhold-Hurek & Hurek, 2011).

Vertical transmission occurs when the endophytic bacteria present in the plant seed colonize the seedling and, further on, the adult plant (Santoyo et al., 2016).

The colonization by soil bacteria into the root system is described as being related to the production of root exudates and rhizodeposits by the plant host (Hardoim et al., 2015). These exudates are the product of photosynthetic processes and comprise low (amino acids, sugars, organic acids, phenolics) and high molecular weight compounds (polysaccharides, proteins). The exudates produced are dependent on host species and growth stage, and environmental factors. Some root exudates can act as attractors to bacteria that can then migrate to the inner tissues of the root system (Glick, 2015).

Once inside the inner tissues, endophytic bacteria reside in intercellular spaces (e.g., Gyaneshwar et al., 2001; reviewed in Hardoim et al., 2015). As a result of the alternative pathways of colonization for the different compartments of a plant, there are often divergences in the concentration of bacteria in BG tissues (usually more colonized) and AG tissues. This is often



observed in studies that regard endophytic bacteria where a distinction is made in these compartments (e.g., Wang et al., 2016). The divergences in communities observed for different plant tissues may also be due to tissue-specific chemical and/or physical properties (Hardoim et al., 2015). It is widely accepted that the bacteria that colonize the inner root system are able to migrate to other tissues, namely to AG tissues, and that this dissemination occurs through the vascular system from the roots and shoots (Hardoim et al., 2015; Vandenkoornhuysen et al., 2015; Reinhold-Hurek & Hurek, 2011 and references therein). These migration phenomena are representative of a microbial endosphere continuity (Vandenkoornhuysen et al., 2015 and references therein).

### Diversity of endophytic bacteria

The bacterial taxonomic diversity and abundance present in the endosphere of a plant depends on many factors, including (i) plant-related factors such as the species, sampling tissue, health and growth stage, (ii) soil-related factors such as pH and moisture content, (iii) and abiotic factors such as temperature and altitude (Hardoim et al., 2008; Hardoim et al., 2015 and references therein). The most common approach to assess the diversity of the bacteria present in a sample relies on the sequencing of the 16S rRNA gene. This ca. 1.5 kb gene is ubiquitous in bacteria, contains highly conserved regions and nine variable regions, where the sequence changes reflect a measure of time or evolution, allowing the gene sequence to act as a molecular clock. Although its effectiveness depends on the taxa at study, the properties of 16S rRNA gene make its sequencing and analysis the standard for bacteria identification usually up to the genus level (Janda & Abbott, 2007 and references therein).

The diversity of the bacterial endophytome has been described as mostly pertaining to the phyla Proteobacteria (mainly to the classes Alphaproteobacteria and Gammaproteobacteria), Bacteroidetes, Firmicutes and Actinobacteria (Hardoim et al., 2015). The phyla Planctomycetes, Verrucomicrobia and Acidobacteria are described as being less abundant in the endosphere (Santoyo et al., 2016). Although only a small number of bacterial phyla are the most associated with the endophytic lifestyle, these four phyla contain many genera and species (Hardoim et al., 2015). In 2002, Lodewyckx et al. listed in review over eighty genera of endophytic isolates belonging to both Gram-positive and -negative bacteria. The list of bacterial genera that exhibit the endophytic lifestyle has been ever-growing as a result of advances in culture media, and also in new sequencing technologies that do not rely on culturability of the bacteria. Although the abovementioned four

phyla remain dominant across the endophytome of many plant species, the dominant families and lower-tier taxa are too host-specific or environment-specific to list.

Hydrocarbon-degrading bacteria from the endosphere of *H. portulacoides* have been assessed using a culture-dependent method. This community presented a dominance of Gram-negative bacteria, which included the genera *Pseudomonas*, *Sphingobium*, *Acinetobacter* and *Ochrobactrum*. The Gram-positive genera *Microbacterium* and *Micrococcus* were also dominant (Olivera et al., 2014a). Another study on the culturable fraction of *H. portulacoides*-associated bacteria, namely from the rhizosphere, also used a hydrocarbon-enrichment step, and showed a dominance of Proteobacteria, Actinobacteria and Bacteroidetes (Oliveira et al., 2014b). These assays where enrichment medium is used direct the focus of the study to a subset of the community, not assessing the whole diversity of the halophyte's endophytome.

#### Plant growth promoting traits

Endophytic bacteria can present PGP abilities by direct and indirect mechanisms. Direct mechanisms include modulating phytohormone levels and enhancing nutrient uptake, while indirect mechanisms are related to competition with phytopathogens, thus decreasing the incidence and damages of diseases (Glick, 2015; Santoyo et al., 2016). Other mechanisms of inducing plant growth are known, such as the production of adenine ribosides, volatile compounds and polyamines, but are less studied (Hardoim et al., 2015 and references therein). The interactions between plant host and beneficial bacteria are highly dependent on (i) plant species or even cultivar, and plant growth stage, (ii) the species or even strain of the bacteria, (iii) biotic factors such as the presence of phytopathogens, (iv) and abiotic factors such as soil properties, presence of contaminants, and temperature (Glick, 2015).

Nutrition enhancement methods involve nutrients such as nitrogen, phosphorus and iron (Glick, 2015; Santoyo et al., 2016). Nitrogen is required in molecules such as nucleic acids and proteins, and can be fixed as atmospheric nitrogen by diazotrophic bacteria but not directly by plants (Glick, 2015). The association with diazotrophic bacteria is especially favorable for the plant host in nitrogen-poor environments (Hardoim et al., 2015). Phosphorus is required as a component of cell membranes and for the synthesis of nucleic acids. Most of the phosphorus in the soil is, however, not available for uptake by the plant. Some beneficial bacteria are able to secrete compounds that allow for the solubilization of phosphorous making it available for plant host uptake (Glick, 2015).

Iron is essential for living organisms as a cofactor of proteins involved in several cellular processes. The available iron in the soil is mostly present as insoluble ferric hydroxides which cannot be easily assimilated by either plant or bacterial cells. Consequently, iron sequestration mechanisms are necessary and occur in bacteria, fungi and some plants via the production and excretion of low molecular weight molecules called siderophores. These molecules are able to bind ferric iron, and this siderophore-iron complex can then be recognized by specific membrane receptors and enter the cell. Bacteria that produce siderophores are then able to supply iron for the nutritional needs of the plant, directly influencing the growth of the host (Glick, 2015).

Direct PGP mechanisms include the production of phytohormones from different classes such as auxins, cytokinins and gibberellins, and participating in the decrease of the stress-related phytohormone ethylene (Glick, 2015; Santoyo et al., 2016). Auxins are known to influence plant processes such as root and shoot growth, tissue differentiation and cell division. Indol-3-acetic acid (IAA) is an auxin that has been found to be produced by ca. 85 % of rhizospheric bacteria (Glick, 2015) and is thought to affect host growth by interfering with the plant defense system (Hardoim et al., 2015). Phytohormones belonging to the class cytokinins are known to promote plant cell division, seed germination, root elongation, differentiation of xylem and chloroplasts, fruit and flower development, nutritional signaling, leaf senescence and plant-pathogen interactions (Glick, 2015). This class of phytohormone is commonly observed in endophytes (Hardoim et al., 2015). Gibberellins are known to increase stem growth, flowering, and leaf and fruit senescence (Glick, 2015). The compound 1-aminocyclopropane-1-carboxylate (ACC) is the precursor for the production of the stress phytohormone ethylene. In the presence of biotic or abiotic stresses, plants can produce an excess of the hormone ethylene which inhibits plant growth. Beneficial bacteria that reduce ethylene levels in the plant do so by producing the enzyme ACC deaminase which cleaves the ACC molecule (reviewed in Santoyo et al., 2016; Glick, 2015; Hardoim et al., 2015).

Indirect PGP mechanisms are mainly those that allow the endophytic bacteria to compete with phytopathogenic organisms. Beneficial bacteria that present these mechanisms limit the plant growth inhibition imposed by phytopathogens. The competition can be performed by means of enzymes that attack cell walls of phytopatogens, production of other antimicrobials, or competing for the same habitat and nutrients (Santoyo et al., 2016; Glick, 2015; Hardoim et al., 2015). One of the methods to compete for nutrition relies on siderophore production, since beneficial bacteria are sequestering iron from the environment, making it unavailable for phytopathogens (Glick, 2015; Hardoim et al., 2015). Bacteria can also be beneficial in the fight against phytopathogens by

producing cell wall degrading enzymes such as chitinases, proteases or lipases, which are able to lyse fungal cells (Glick, 2015). Endophytic bacteria can also lead to a higher tolerance to phytopathogens by inducing plant defense reactions in an initial phase of colonization. This trigger in plant immune response allows the host to engage against pathogens in a more efficient manner (Hardoim et al., 2015 and references therein).

As a consequence of direct and indirect PGP mechanisms, the plant fitness is enhanced resulting in the perseverance of the plant host in the presence of biotic and abiotic stresses. The mechanisms employed by PGP endophytic bacteria can amount to several outcomes such as enhanced seed germination, increased growth of seedlings, roots and shoots, increased biomass, improved tolerance to biotic and abiotic stresses, production of secondary metabolites, and better plant nutrition (Glick, 2015). As an example of PGP endophytes in a closely related plant species, a study conducted on the halophyte *Arthrocnemum macrostachyum* (same family as *H. portulacoides*), showed that isolates from its endosphere exhibited PGP abilities. Eight isolates belonging to the family *Bacillaceae* exhibited PGP traits *in vitro* in the presence of salt, including enzymatic activities (pectinolytic, chitinolytic, cellulolytic, amylolytic and proteolytic), IAA production, phosphate solubilization and siderophore production. A consortium of promising isolates was tested for their PGP abilities *in vivo*, and was able to improve seed germination and enhance the salt tolerance of the plant host (Navarro-Torre et al., 2016).

The endosphere is a relevant source of bacteria with extensive diversity. This diversity has been studied using different approaches, from classic microbiology by plating out macerated surface-sterilized plant tissues, to molecular techniques. More recent advances in sequencing technologies have made available a plethora of molecular-based techniques: clone library construction, denaturing gradient gel electrophoresis profile analysis, 454-pyrosequencing, and most recently Illumina sequencing. In recent years, the most used approaches have been based on 16S rRNA gene sequencing from complex communities, however, whole genome metagenomics of the communities has been emerging recently (e.g., Akinsanya et al., 2015). Understanding the core diversity present in the endosphere of a plant allows for the development of plant-specific growth-promoting agents, which can be of chief importance for downstream applications such as remediation or detoxification processes.

## 1.4 References

- Akinsanya, M. A., Goh, J. K., Lim, S. P. & Ting, A. S. Y. (2015). Metagenomics study of endophytic bacteria in Aloe vera using next-generation technology. *Genomics Data* **6**, 159-163.
- Almeida, C. M. R., Mucha, A. P., Delgado, M. F. C., Caçador, M. I., Bordalo, A. A. & Vasconcelos, M. T. S. D. (2008). Can PAHs influence Cu accumulation by salt marsh plants? *Mar Environ Res* **66**, 311-318.
- Amari, T., Ghnaya, T. & Abdelly, C. (2017). Nickel, cadmium and lead phytotoxicity and potential of halophytic plants in heavy metal extraction. *S Afr J Bot* **111**, 99-110.
- Andreotti, F., Mucha, A. P., Caetano, C., Rodrigues, P., Gomes, C. R. & Almeida, C. M. R. (2015). Interactions between salt marsh plants and Cu nanoparticles (NPs) – Effects on metal uptake and phytoremediation processes. *Ecotoxicol Environ Saf* **120**, 303-309.
- Anjum, N. A., Ahmad, I., Válega, M., Pacheco, M., Figueira, E., Duarte, A. C. & Pereira, E. (2011). Impact of seasonal fluctuations on the sediment-mercury, its accumulation and partitioning in *Halimione portulacoides* and *Juncus maritimus* collected from Ria de Aveiro Coastal Lagoon (Portugal). *Water Air Soil Pollut* **222**, 1-15.
- Anjum, N. A., Israr, M., Duarte, A. C., Pereira, M. E. & Ahmad, I. (2014). *Halimione portulacoides* (L.) physiological/biochemical characterization for its adaptive responses to environmental mercury exposure. *Environ Res* **131**, 39-49.
- Bouchard, V., Creach, V., Lefeuvre, J. C., Bertru, G. & Mariotti, A. (1998). Fate of plant detritus in a European salt marsh dominated by *Atriplex portulacoides* (L.) Aellen. *Hydrobiologia* **373**, 75-87.
- Caçador, I., Caetano, M., Duarte, B. & Vale, C. (2009). Stock and losses of trace metals from salt marsh plants. *Mar Environ Res* **67**, 75-82.
- Canário, J., Caetano, M., Vale, C. & Cesário, R. (2007). Evidence for elevated production of methylmercury in salt marshes. *Environ Sci Technol* **41**, 7376-7382.
- Canário, J., Poissant, L., Pilote, M., Caetano, M., Hintelmann, H. & O'Driscoll, N. J. (2017). Salt-marsh plants as potential sources of Hg<sup>0</sup> into the atmosphere. *Atmos Environ* **152**, 458-464.
- Carvalho, P. N., Basto, M. C. P., Silva, M. F. G. M., Machado, A., Bordalo, A. A. & Vasconcelos, M. T. S. D. (2010). Ability of salt marsh plants for tributyltin (TBT) remediation in sediments. *Environ Sci Pollut Res* **17**, 1279-1286.
- Costa, C. & Jesus-Rydin, C. (2001). Site investigation on heavy metals contaminated ground in Estarreja - Portugal. *Eng Geol* **60**, 39-47.
- Couto, M. N. P. F. S., Basto, M. C. P. & Vasconcelos, M. T. S. D. (2011). Suitability of different salt marsh plants for petroleum hydrocarbons (PHC) remediation. *Chemosphere* **84**, 1052-1057.

- Everest, A. & Ozturk, E. (2005).** Focusing on the ethnobotanical uses of plants in Mersin and Adana provinces (Turkey). *J Ethnobiol Ethnomed* **6**, 1-6.
- Fernandes, J. P., Almeida, C. M. R., Andreotti, F., Barros, L., Almeida, T. & Mucha, A. P. (2016).** Response of microbial communities colonizing salt marsh plants rhizosphere to copper oxide nanoparticles contamination and its implications for phytoremediation processes. *Sci Total Environ* **581-582**, 801-810.
- Figueira, E., Freitas, R., Pereira, E. & Duarte, A. (2012).** Mercury uptake and allocation in *Juncus maritimus*: implications for phytoremediation and restoration of a mercury contaminated salt marsh. *J Environ Monit* **14**, 2181-2188.
- Glick, B. R. (2015).** Beneficial plant-bacterial interactions. Springer, Heidelberg.
- Gyaneshwar, P., James, E. K., Mathan, N., Reddy, P. M., Reinhold-Hurek, B. & Ladha, J. K. (2001).** Endophytic colonization of rice by a diazotrophic strain of *Serratia marcescens*. *J Bacteriol* **183**, 2634-2645.
- Hardoim, P. R., van Overbeek, L. S. & van Elsas, J. D. (2008).** Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* **16**, 463-471.
- Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., Döring, M. & Sessitsch, A. (2015).** The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol Mol Biol Rev* **79**, 293-320.
- Henning K., Villforth F. (1940).** Experimentelle untersuchungen zur frage der bacteriesymbiose in höheren pflanzen und ihre beeinflussung durch 'Leitemente'. *Biochem Z* **305**, 299-309.
- Janda, J. M. & Abbott, S. L. (2007).** 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* **45**, 2761-2764.
- Lodewyckx, C., Vangronsveld, J., Porteous, F., Moore, E. R. B., Taghavi, S., Mezgeay, M. & van der Lelie, D. (2002).** Endophytic bacteria and their potential applications. *Crit Rev Plant Sci* **21**, 583-606.
- Lopes, J. F., Dias, J. M. & Dekeyser, I. (2001).** Influence of tides and river inputs on suspended sediment transport in the Ria de Aveiro lagoon, Portugal. *Phys Chem Earth* **26**, 729-734.
- Lozano-Rodriguez, E., Hernández, L. E., Bonay, P. & Carpena-Ruiz, R. O. (1997).** Distribution of cadmium in shoot and root tissues of maize and pea plants: physiological disturbances. *J Exp Bot* **48**, 123-128.
- Navarro-Torre, S., Barcia-Piedras, J. M., Mateos-Naranjo, E., Redondo-Gómez, S., Camacho, M., Caviedes, M. A., Pajuelo, E. & Rodríguez-Llorente, I. D. (2016).** Assessing the role of endophytic bacteria in the halophyte *Arthrocnemum macrostachyum* salt tolerance. *Plant Biol (Stuttg)* **19**, 249-256.
- Negrin, V. L., Teixeira, B., Godinho, R. M., Mendes, R. & Vale, C. (2017).** Phytochelatins and monothiol in salt marsh plants and their relation to metal tolerance. *Mar Pollut Bull* doi: 10.1016/j.marpolbul.2017.05.045.

- Oliveira, V., Gomes, N. C. M., Almeida, A., Silva, A. M. S., Simões, M. Q. M., Smalla, K. & Cunha, A. (2014a).** Hydrocarbon contamination and plant species determine the phylogenetic and functional diversity of endophytic degrading bacteria. *Mol Ecol* **23**, 1392-1404.
- Oliveira, V., Gomes, N. C. M., Cleary, D. F. R., Almeida, A., Silva, A. M. S., Simões, M. M. Q., Silva, H. & Cunha, A. (2014b).** Halophyte plant colonization as a driver of the composition of bacterial communities in salt marshes chronically exposed to oil hydrocarbons. *FEMS Microbiol Ecol* **90**, 647-662.
- Partida-Martínez, L. & Heil, M. (2011).** The microbe-free plant: fact or artifact? *Front Plant Sci* **2**, article 100. doi: 10.3389/fpls.2011.00100.
- Pato, P., Lopes, C., Válega, M., Lillebø, A. I., Dias, J. M., Pereira, E. & Duarte A. C. (2008a).** Mercury fluxes between an impacted coastal lagoon and the Atlantic Ocean. *Estuarine, Coastal Shelf Sci* **76**, 787-796.
- Pato, P., Válega, M., Pereira, E., Vale, C. & Duarte, A. C. (2008b).** Inputs from a mercury-contaminated lagoon: impact on the nearshore waters of the Atlantic Ocean. *J Coastal Res* **24**, 28-38.
- Pereira, M. E., Duarte, A. C., Millward, G. E., Abreu, S. N. & Vale, C. (1998a).** An estimation of industrial mercury stored in sediments of a confined area of the Lagoon of Aveiro (Portugal). *Wat Sci Tech* **37**, 125-130.
- Pereira, M. E., Duarte, A. C., Millward, G. E., Vale, C. & Abreu, S. N. (1998b).** Tidal export of particulate mercury from the most contaminated area of Aveiro's Lagoon, Portugal. *Sci Total Environ* **213**, 157-163.
- Pereira, M. E., Lillebø, A. I., Pato, P., Válega, M., Coelho, J. P., Lopes, C. B., Rodrigues, S., Cachada, A., Otero, M., Pardal, M. A. & Duarte, A. C. (2009).** Mercury pollution in Ria de Aveiro (Portugal): a review of the system assessment. *Environ Monit Assess* **155**, 39-49.
- Perotti, R. (1926).** On the limits of biological inquiry on soil science. *Proc Int Soc Soil Sci* **2**, 146-161.
- Quispel, A. (1992).** A search of signals in endophytic microorganisms. In: Verma DPS (ed) *Molecular signals in plant-microbe communications*. CRC, Florida, pp. 471-491.
- Ramalhosa, E., Segade, S. R., Pereira, E., Vale, C. & Duarte A. (2006).** Mercury cycling between the water column and surficial sediments in a contaminated area. *Water Res* **40**, 2893-2900.
- Reinhold-Hurek, B. & Hurek, T. (2011).** Living inside plants: bacterial endophytes. *Curr Opin Plant Biol* **14**, 435-443.
- Santoyo, G., Moreno-Hagelsieb, G., del Carmen Orozco-Mosqueda, M. & Glick, B. R. (2016).** Plant growth-promoting bacterial endophytes. *Microbiol Res* **183**, 92-99.
- The International Plant Names Index (2012).** Published on the Internet <http://www.ipni.org> [accessed 31 January 2017].

- Válega, M., Lillebø, A. I., Pereira, M. E., Caçador, I., Duarte, A. C. & Pardal, M. A. (2008a).** Mercury in salt marshes ecosystems: *Halimione portulacoides* as biomonitor. *Chemosphere* **73**, 1224-1229.
- Válega, M., Lillebø, A. I., Caçador, I., Pereira, M. E., Duarte, A. C. & Pardal, M. A. (2008b).** Mercury mobility in a salt marsh colonized by *Halimione portulacoides*. *Chemosphere* **72**, 1607-1613.
- Válega, M., Lillebø, A. I., Pereira, M. E., Corns, W. T., Stockwell, P. B., Duarte, A. C. & Pardal, M. A. (2008c).** Assessment of methylmercury production in a temperate salt marsh (Ria de Aveiro Lagoon, Portugal). *Mar Pollut Bull* **56**, 136-162.
- Válega, M., Lima, A. I. G., Figueira, E. M. A. P., Pereira, E., Pardal, M. A. & Duarte, A. C. (2009).** Mercury intracellular partitioning and chelation in a salt marsh plant, *Halimione portulacoides* (L.) Aellen: Strategies underlying tolerance in environmental exposure. *Chemosphere* **74**, 530-536.
- Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. (2015).** The importance of the microbiome of the plant holobiont. *New Phytol* **206**, 1196-1206.
- Wang, W., Zhai, Y., Cao, L., Tan, H. & Zhang, R. (2016).** Endophytic bacterial and fungal microbiota in sprouts, roots and stems of rice (*Oryza sativa* L.). *Microbiol Res* **188-189**, 1-8.
- Williams, T. P., Bubbs, J. M. & Lester, J. N. (1994).** Metal accumulation within salt marsh environments: a review. *Mar Pollut Bull* **28**, 277-290.



## Chapter 2

---

### Objectives

## **Contents**

### **2.1 Scope**

### **2.2 Main goal**

### **2.3 Specific goals**

### **2.4 Hypotheses**

## 2.1 Scope

The present work will explore the endophytic bacterial community of the salt marsh plant *Halimione portulacoides*. For this, three main aspects are considered for the scope of the document:

I. The halophyte *H. portulacoides* presents phytoremediation capability, and its endosphere is an unexplored potential source of bacteria that promote plant growth and, subsequently, enhance phytoremediation processes. Endophytic bacteria will be obtained from sites with different degrees of metal(loid) contamination, using diverse culture media so as to obtain a representative collection. The isolates will then be subjected to several tests aiming at a deep characterization of the collection in regards to taxonomy, enzymatic activity and plant growth promotion (PGP) traits. Promising candidates for PGP will be further tested for their ability to promote growth in model plants.

II. The endosphere is a relatively unexplored environment. As a consequence, there is a high likelihood that the endophytic bacteria collection will contain novel bacterial taxa, whose description will represent incremental advances in the field of microbiology. For this purpose, representatives of novel taxa will be described and added to databases respecting established regulations.

III. The in-depth exploration of the endosphere by culture-independent methods has always faced the obstacle of host DNA contamination. As such, finding a way to overcome this problem is fundamental for this type of studies. In-depth analysis of the diversity in the endosphere will be assessed using PCR blockers that avoid host DNA contamination.

This work, overall, has the potential to provide in-depth knowledge on bacterial diversity of endophytes, on bacterial strains that are adequate for promotion of halophyte growth in contaminated sites, and on novel bacterial species.

## 2.2 Main goal

The main goal of the present work is to characterize endophytic bacterial communities associated with the salt marsh plant *Halimione portulacoides*.

## 2.3 Specific goals

In order to accomplish the proposed main goal, five specific goals were established:

- I. Establish a collection of endophytic bacteria from *H. portulacoides* sampled from sites with high, intermediate and no metal(loid) contamination.
- II. Understand the structure, diversity and properties of the collection of endophytic bacteria.
- III. Identify potential candidates for *in vivo* PGP and perform tests on model plants.
- IV. Identify and describe novel taxa within the endophytic bacteria collection.
- V. Gain a deeper insight into the structure and diversity of the endophytic bacteria associated with *H. portulacoides* by combining information from culture-based and culture-independent approaches.

## 2.4 Hypotheses

The following are hypotheses to be tested throughout the work:

- I. The endosphere of the halophyte *H. portulacoides* is a hotspot for bacterial diversity.
- II. The presence of metal(loid)s influences the structure and diversity of the endophytic bacteria of *H. portulacoides*.
- III. The endophytic bacterial community of the halophyte includes bacteria that have potential for promoting plant growth.
- IV. The endophytic bacterial community of the halophyte includes a plethora of novel bacteria taxa.

## Chapter 3

---

Culturable endophytic bacteria from *Halimione portulacoides*

## Contents

**3.1** Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization, and influence of metal(loid) contamination

**3.2** Diversity of endophytic *Pseudomonas* in *Halimione portulacoides* from metal(loid)-polluted salt marshes

### **3.1 Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization, and influence of metal(loid) contamination**

#### **Authors**

Cátia Fidalgo<sup>1,2</sup>, Isabel Henriques<sup>2</sup>, Jaqueline Rocha<sup>1</sup>, Marta Tação<sup>1</sup>, Artur Alves<sup>1</sup>

<sup>1</sup> CESAM, Departamento de Biologia, Universidade de Aveiro, Aveiro, Portugal

<sup>2</sup> iBiMED and CESAM, Departamento de Biologia, Campus de Santiago, Universidade de Aveiro, 3810-193 Aveiro, Portugal

#### **Publication status**

Published

Environ Sci Pollut Res

2016 May; 23(10):10200-14

doi: 10.1007/s11356-016-6208-1

## Abstract

*Halimione portulacoides* is abundant in salt marshes, accumulates mercury (Hg), and was proposed as useful for phytoremediation and pollution biomonitoring. Endophytic bacteria promote plant growth and provide compounds with industrial applications. Nevertheless, information about endophytic bacteria from *H. portulacoides* is scarce. Endophytic isolates ( $n = 665$ ) were obtained from aboveground and belowground plant tissues, from two Hg-contaminated sites (sites E and B) and a non-contaminated site (site C), in the estuary Ria de Aveiro. Representative isolates ( $n = 467$ ) were identified by 16S rRNA gene sequencing and subjected to functional assays. Isolates affiliated with Proteobacteria (64 %), Actinobacteria (23 %), Firmicutes (10 %), and Bacteroidetes (3 %). *Altererythrobacter* (7.4 %), *Marinilactibacillus* (6.4 %), *Microbacterium* (10.2 %), *Salinicola* (8.8 %), and *Vibrio* (7.8 %) were the most abundant genera. Notably, *Salinicola* ( $n = 58$ ) were only isolated from site C; *Hoeflea* (17), *Labrenzia* (22), and *Microbacterium* (67) only from belowground tissues. This is the first report of *Marinilactibacillus* in the endosphere. Principal coordinate analysis showed that community composition changes with the contamination gradient and tissue. Our results suggest that the endosphere of *H. portulacoides* represents a diverse bacterial hotspot including putative novel species. Many isolates, particularly those affiliated to *Altererythrobacter*, *Marinilactibacillus*, *Microbacterium*, and *Vibrio*, tested positive for enzymatic activities and plant growth promoters, exposing *H. portulacoides* as a source of bacteria and compounds with biotechnological applications.

## Keywords

Endophyte, bacteria, *Halimione portulacoides*, salt marsh plants, plant growth promotion, extracellular enzymes



## Background

Salt marshes are among the most productive ecosystems in terms of primary production and are a habitat to several species with commercial value (Alongi, 1998). The vegetation in this type of ecosystem is usually subjected to harsh environmental conditions including periodic tidal flooding, high salinity levels, and contaminants such as metals from urban and industrial effluents. As a consequence, the typical salt marsh vegetation presents low species diversity and high specialization (Woerner and Hackney, 1997).

Ria de Aveiro, a costal lagoon on the northwest of Portugal, is one of the most mercury-contaminated coastal systems in Portugal. For over four decades (1950–1994), industrial effluents containing mercury (Hg), arsenic (As), zinc (Zn), lead (Pb), and other contaminants were discharged into natural water streams of this lagoon which led to contamination of water and sediments of adjacent marshes (Costa and Jesus-Rydin, 2001). Although discharges have substantially decreased over the last two decades, sediments in the Laranjo Bay salt marshes are still considerably contaminated, namely with Hg. This highly toxic nonessential metal is a hazardous contaminant with potential to be released to the water column and assimilated by the biota (Pereira et al., 2009).

The perennial halophyte *Halimione portulacoides* (L.) Aellen is one of the most productive and abundant species in European salt marshes (Bouchard et al., 1998), including those along the Ria de Aveiro lagoon. This Hg-tolerant halophyte allocates this toxic metal preferably to belowground tissues and was proposed as a bioindicator and biomonitor for Hg pollution (Válega et al., 2008a). Additionally, this halophyte has been proposed for phytoremediation of Hg (Anjum et al., 2011), tributyltin (Carvalho et al., 2010), and petroleum hydrocarbons (Couto et al., 2011).

Endophytic bacteria colonize internal plant tissues without causing symptoms of disease (Lodewyckx et al., 2002). These bacteria are highly diverse taxonomically, ubiquitous from herbaceous plants to trees, and contribute to plant health and productivity (Berg et al., 2014). Endophytes present a great potential for applications in agricultural, medical, biotechnological, and industrial fields. Industrially important enzymes such as amylases, cellulases, lipases, pectinases, proteases, and xylanases, among others, have previously been found in endophytic bacteria (e.g., Lodewyckx et al., 2002).

Plant-associated bacteria, including endophytic bacteria, are known to be very relevant for plant growth promotion (PGP; Hardoim et al. 2008), including in cases of biotic or abiotic stresses, such as contaminated soils (reviewed in Rajkumar et al., 2012). This PGP action may be conducted

through direct or indirect mechanisms. Direct PGP mechanisms include enhancement of plant nutrient acquisition, by aiding in nitrogen fixation and phosphate solubilization; plant growth by production of the auxin indol-3-acetic acid (IAA); and lowering of ethylene-related plant stress by producing the enzyme 1-aminocyclopropane-1-carboxylate (ACC) deaminase. Indirect PGP is achieved by preventing phytopathogen prejudicial effects on plant development by producing metabolites involved in antimicrobial activities and iron chelation (reviewed in Bulgarelli et al., 2013). It is also known that plant-associated bacteria are able to aid in phytoremediation processes by phytoextraction (reviewed in Sessitsch et al., 2013). In fact, Oliveira et al. (2014a) have shown that endophytic bacteria isolated from *H. portulacoides* have potential to be used in detoxification of petroleum hydrocarbons.

Notwithstanding the acknowledged importance of endophytic bacteria, comprehensive studies regarding the basal culturable communities of these bacteria in the specific environment of salt marshes are scarce, as studies are usually directed toward studying contaminant-degrading bacteria. *Halimione portulacoides* is a highly relevant and ecologically important halophyte in salt marshes; however, the structure of endophytic bacterial communities of this plant has not yet been fully explored, only its fraction that is able to degrade hydrocarbons (Oliveira et al., 2014a).

This study aimed to characterize the culturable endophytic bacterial communities of this halophyte from a highly contaminated site, an intermediately contaminated site, and a non-contaminated site in Ria de Aveiro, regarding their phylogenetic diversity and traits associated with direct and indirect PGP effects.

## Methods

### Study Area and Sampling

The study was conducted in the estuary Ria de Aveiro, in Aveiro, Portugal. A confined area in this lagoon was subjected to Hg-rich industrial discharges for four decades (1950s–1990s, Pereira et al., 1998). Three locations in the salt marsh of Ria de Aveiro were assessed: one highly contaminated by industrial activity (site B, 450 m downstream from the contamination source in the estuary, 40° 43' 47.46" N, 8° 36' 45.40" W), one with intermediate levels of contamination (site E, 2500 m downstream from the contamination source, 40° 43' 15.90" N, 8° 38' 15.76" W) (Válega et al., 2008b), and a non-contaminated location (site C, 40° 38' 05.36" N, 8° 39' 38.87" W; 11.3 km from the contamination source). Healthy specimens of the halophyte *H. portulacoides* were sampled in monospecific stands, at least in triplicate in each site. Adjacent sediment was sampled in triplicate with a 3 cm in diameter and 13 cm depth corer. Sampling was performed during low tide in November 2012 (average temperature 16 °C). Samples were transported to the laboratory under refrigerated conditions and promptly processed.

### Metal(loid) quantification

Composite samples of aboveground (AG) tissues (leaves, stems, and flowers; 1 g fresh weight), belowground (BG) tissues (roots; 1 g FW), and sediment (2 g FW) from each site were used for metal(loid) quantification, in triplicate. Chromium (Cr), nickel (Ni), copper (Cu), zinc (Zn), As (as total arsenic), and mercury (Hg) were quantified. Samples were subjected to acid digestion with concentrated nitric acid overnight at 115 °C (Figueira and Freitas, 2013) and analyzed by inductively coupled plasma mass spectrometry (ICP-MS) according to the International Standard ISO 17294 by a certified laboratory at the University of Aveiro.

One-way ANOVA (function *aov*) combined with post hoc Tukey's honest significant differences method (function *TukeyHSD*) were used to assess differences in metal(loid) concentrations in sediment and plant tissue samples on all sites. Welch two-sample *t* test (function *t.test*) was used to estimate differences in metal(loid) concentrations (i) among plant tissues, for each isolation site and (ii) between sediment and plant samples, for each isolation site. These analyses were performed using the *stats* package in the free software R (R Core Team, 2014) with abovementioned functions *aov*, *TukeyHSD*, and *t.test*, considering a 95 % confidence interval.

### **Isolation of endophytic bacteria, growth media, and conditions**

Eleven plants were processed at the laboratory: five from site C, three from site B, and three from site E. First, all plants were thoroughly washed with tap water. For each plant, 5 g (FW) of AG tissues and 2 g (FW) of BG tissues were used for endophytic bacteria isolation. Surface sterilization was optimized and performed as follows: samples were sequentially immersed in 50 mL of phosphate-buffered saline (PBS; 10 min), 96 % ethanol (1 min), 5 % NaOCl (30 min), 96 % ethanol (1 min), and rinsing in distilled sterile water three times. Surface sterilization was confirmed by inoculating water from the last wash in Tryptic Soy Agar (TSA, Merck, Germany), Marine Agar (MA, Difco Laboratories, France), and R2A (Merck, Germany) culture media. The plates were examined for bacterial growth after incubation at 28 °C for 72 h. Tissues were then ground with a pestle in a mortar containing 10 mL of PBS. Serial dilutions were prepared in PBS and spread, in duplicate, on TSA, MA, and R2A media. The plates were incubated up to 72 h at 28 °C and were observed every 24 h for colony-forming unit (CFU) count. Morphologically distinct endophytic bacterial colonies were selected and purified by streaking, resulting in a collection of 665 endophytic bacterial isolates, and stored at -80 °C in 20 % glycerol. Function *aov* combined with *TukeyHSD* was used to assess differences in average CFU per gram of tissue (fresh weight) in different isolation sites and culture media, and *t.test* was used to assess tissue-based differences.

### **DNA extraction, PCR-based fingerprinting, and selection of representative isolates**

DNA from all 665 isolates was extracted using the Genomic DNA Purification kit #0513 (Thermo Scientific, USA) according to the manufacturer's instructions. BOX-PCR with BOXA1 primer (5'-CTACGGCAAGGCGACGCTGAC-3'; Versalovic et al., 1994) was used for molecular typing of all isolates. PCR reactions were performed in a volume of 25 µL, using 5× Green GoTaq buffer, GoTaq DNA polymerase (5U µL<sup>-1</sup>), MgCl<sub>2</sub> (25 mM), and a dNTP mixture (2 mM) (Promega, USA). Conditions for amplification were as follows: one cycle at 95 °C (7 min), 30 cycles at 94 °C (1 min), 53 °C (1 min), 65 °C (8 min), and a final cycle at 65 °C (16 min). One hundred thirty-three isolates did not yield a clear profile with this methodology and were typed using ERIC-PCR with ERIC1R (5'-ATGTAAGCTCCTGGGGATTAC-3'; Versalovic et al., 1991) and ERIC2 (5'-AAGTAAGTGACTGGGGTGAGCG-3'; Versalovic et al., 1991) primers. Conditions for PCR amplification were the same as for BOX-PCR, except the annealing temperature, which was 52 °C.

Analysis of the genetic fingerprinting patterns was performed with GelCompar II software (Applied Maths, Belgium). The Pearson correlation coefficient was applied, and cluster analysis was performed using the unweighted pair group method with arithmetic mean (UPGMA) algorithm. The resulting dendrograms were analyzed in order to obtain groups of isolates with at least 85 % similarity. This cutoff was determined so that patterns that were known to be equal (molecular-weight size markers) would be considered to be in the same group. Representative isolates of each group were randomly selected. A total of 467 representative isolates were obtained: 419 from molecular typing and 48 isolates that did not yield clear profiles by either BOX- or ERIC-PCR.

### **16S rRNA gene amplification, phylogenetic affiliation, and multivariate analyses**

Four hundred sixty-seven representative isolates were subjected to PCR amplification of 16S rRNA gene using the universal primers 27f (5'-AGAGTTTGATCCTGGCTCAG-3'; Lane, 1991) and 1492r (5'-GGTACCTTGTACGACTT-3'; Lane, 1991) and the NZYtaq 2× Green Master Mix (NZYTech, Portugal) in 25-μL reactions: one cycle at 94 °C (5 min), 30 cycles at 94 °C (1 min), 55 °C (1 min), 72 °C (1.5 min), and a final cycle at 72 °C (10 min). Five isolates did not yield quality sequences after several attempts and were excluded from the analysis. The remaining 462 amplification products were purified with the DNA Clean & Concentrator™-5 (Zymoresearch, USA) according to manufacturer's instructions. Amplicons were partially sequenced with primer 27f using GATC Biotech services (Germany) according to company specifications. The partial 16S rRNA gene sequences obtained in this study were submitted to GenBank and are available under the accession numbers KT324749 to KT325209. Quality sequences with an overall average of 893 bp were examined and edited with FinchTV V1.4.0 software (Geospiza, USA). The Basic Local Alignment Search Tool (BLAST; Altschul et al., 1997) and the EzTaxon server (Kim et al., 2012) were used to obtain taxonomic classification up to the genus level. Information from the representative isolates was used to associate taxonomic affiliation to the respective 658 isolates from the collection. Sequences were then subjected to the Ribosomal Database Project (RDP) pyrosequencing pipeline (Cole et al., 2014), where they were aligned and clustered into operational taxonomic units (OTUs) considering a distance limit of 0.03 (97 % similarity). Consequently, an OTU table was built with abundances for 21 samples: five for CAG—community from AG tissues from site C; five for CBG—BG tissues from site C; two for EAG—AG tissues from site E; three for EBG—BG tissues from site E; three for BAG—AG tissues from site B; and three for BBG—BG tissues from site B. Three AG samples were collected from site E;

however, only two yielded culturable bacteria. As such, only those were considered for subsequent analyses.

Using R version 3.1.1 (R Core Team, 2014), several analyses were performed to characterize the bacterial communities regarding their OTU abundance. Diversity indices Shannon's  $H'$  diversity and Pielou's evenness were assessed with functions *diversity* and with *diversity/log(specnumber)*, respectively, from the *vegan* package (Oksanen et al., 2015). The *multipatt* function from package *indicspecies* (De Caceres and Legendre, 2009) was used to find indicator species (here, OTUs) in samples from different sites, tissues and communities, using a 10 % significance level for selecting indicators. The species-site group association parameter "IndVal.g" was used, since it corrects for unequal group sizes.

The variation in OTU composition among the samples was analyzed by assessing differences in a dissimilarity matrix, built with Bray-Curtis dissimilarity measure (function *vegdist* from package *vegan* in R) after transforming the OTU abundance table with square root (function *sqrt* in R). PERMANOVA was used to analyze variances in the Bray-Curtis dissimilarity matrix using the *adonis* function (from package *vegan* in R) on the interaction between isolation site and tissue, and to both variables independently. For this analysis, the number of permutations was 999 in all tests.

Information in the Bray-Curtis dissimilarity matrix was then graphically assessed with a principal coordinate analysis (PCoA) using *cmdscale* function (from package *vegan* in R) to create the PCoA ordination. The significance of the observed patterns in the PCoA was analyzed by assessing differences in sample scores across PCoA axis 1 (PCoA1) and PCoA axis 2 (PCoA2). The normality test *shapiro.test* (*stats* package in R) revealed non-normality in PCoA1 (Shapiro-Wilk:  $W=0.8848$ ,  $p$ -value  $< 0.05$ ) and normality in PCoA2 (Shapiro-Wilk:  $W = 0.9352$ ,  $p$ -value  $> 0.1$ ). As a consequence, patterns in PCoA1 were assessed with Kruskal-Wallis rank sum test (*kruskal.test* from *stats* package in R) combined with *posthoc.kruskal.nemenyi.test* (from *PMCMR* package in R; Pohlert 2005) for samples from different sites and communities, and with Wilcoxon rank sum test (*wilcox.test* from *stats* package in R) for samples from different tissues. PCoA2 patterns were analyzed with *aov* for samples from different sites and communities, and with *t.test* for samples from different tissues.

### **Screening of endophytic bacteria for enzymatic activity**

The 467 representative isolates were individually screened, in triplicate, for different enzymatic activities: amylolytic, cellulolytic, lipolytic, pectinolytic (at pH 5.0 and 7.0), proteolytic, and xylanolytic. The ability of the isolates to degrade a substrate was evaluated using 0.2 % starch, 0.5 % carboxymethyl cellulose, 1 % Tween 20, 0.5 % pectin, 1 % skim milk, or 0.5 % xylan, respectively, on TSA or MA culture medium. The basal culture medium (TSA or MA) was selected for each isolate, according to where best bacterial growth was obtained. Tween 20 and skim milk solutions were sterilized separately and only added to the culture media after cooling down. The agar medium plates were inoculated with 5  $\mu$ L of a bacterial suspension with optical density (OD) at 600 nm=0.8 U and incubated at 30 °C until sufficient growth was obtained, or up to 120 h. The enzymatic activities were assessed according to the presence (positive result) or absence (negative result) of a halo surrounding the bacterial growth. Amylolytic activity was detected when a clear halo was observed upon addition of Lugol solution to the agar plate. Cellulolytic and xylanolytic activities were detected when a clear halo was observed upon incubation with a 1 % Congo red solution (with 10 % ethanol) and subsequent wash with 1 M NaCl. Lipolytic activity was observed by the formation of a white precipitate around the bacterial growth. Pectinolytic activity was detected by formation of a clear halo after addition of 1 % hexadecyltrimethylammonium bromide (CTAB) solution. Proteolytic activity was revealed by visualization of a clear halo against an opaque background. The results were interpreted according to the visibility of the halo obtained: negative, no halo observed; positive, faint to intense halo observed.

### **Screening of endophytic bacteria for plant growth promotion traits**

Representative isolates were individually screened for different PGP traits: ACC deaminase activity, IAA production, phosphate solubilization, and siderophore production. Isolates were incubated in adequate growth media at 30 °C until sufficient growth was obtained. Unless otherwise stated, strains for negative and positive controls were studied in Pereira et al. (2013).

To determine ACC deaminase activity, isolates were grown in DF salt minimal medium (Dworkin and Foster, 1958) or variations of this medium. Basal DF was prepared as follows (per liter): 4 g  $\text{KH}_2\text{PO}_4$ , 6 g  $\text{Na}_2\text{HPO}_4$ , 0.2 g  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 2 g glucose, 1 mg  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 10  $\mu\text{g}$   $\text{H}_3\text{BO}_3$ , 10  $\mu\text{g}$   $\text{MnSO}_4$ , 70  $\mu\text{g}$   $\text{ZnSO}_4$ , 50  $\mu\text{g}$   $\text{CuSO}_4$ , 10  $\mu\text{g}$   $\text{MoO}_3$ , 12 g agar, pH 7.2. This medium was either supplemented with 3 mM ACC (DF + ACC; ACC was sterilized by filtration and added to cooled down medium) or 2

g L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (DF + ammonium sulfate) or no nitrogen source (DF). The ability to use ACC as a sole nitrogen source is a consequence of the activity of the enzyme ACC deaminase (Penrose and Glick, 2003). *Pseudomonas putida* EAPC8 and *Arthrobacter rombi* EC32A were used as positive controls for ACC deaminase activity.

To test for IAA production, the method of Gordon and Weber (1951) was used: for each isolate, 500 µL of supernatant obtained from cultures grown in MB or TSB + L-tryptophan (1 %) was mixed with 1 mL of Salper solution. After incubating in the dark for 30 min, OD was read at 535 nm. Using a calibration curve of pure IAA as a standard, IAA concentrations (µg mL<sup>-1</sup>) were obtained for each isolate. *Pseudomonas putida* EAPC8 and *Pseudomonas fluorescens* S3X were used as positive controls for IAA production.

The ability of isolates to solubilize phosphate was screened on National Botanical Research Institute's Phosphate (NBRIP) solid growth medium (Nautiyal, 1999). After inoculation and appropriate incubation, phosphate solubilization was observed as a clear halo surrounding the bacterial growth. *Pseudomonas reactans* EDP28, *P. putida* EAPC8, *A. rombi* EC32A, and *P. fluorescens* S3X were used as negative controls for phosphate solubilization, and *Arthrobacter nicotinovorans* EAPAA and *Rhodococcus* sp. EC35 were used as positive controls.

In order to test for siderophore production, all glass materials used were deferrated by treating with 6 % nitric acid and extensively washing with sterile double distilled water. Culture media deferration was conducted on TSA and MA medium as described by Cox (1994). After inoculation and incubation on either medium, the plates with bacterial growth were overlaid with CAS (O-CAS; Pérez-Miranda et al., 2007) and incubated with this overlay for 2 h. Siderophore production was detected as a change of color of the overlay from blue to orange or purple. This assay was only performed with TSA medium, since the O-CAS reacted with MA medium preventing a clear reading of the result. *Pseudomonas reactans* EDP28, *P. putida* EAPC8, *A. rombi* EC32A, *P. fluorescens* S3X, *A. nicotinovorans* EAPAA, and *Rhodococcus* sp. EC35 were used as positive controls for siderophore production.

The representative isolates were also assessed for their nitrogen fixing abilities by PCR-screening for the presence of *nifH* gene: the primers IKG3 and DVV (Ando et al., 2005), described as the best pair of primers to use in *nifH* amplification (Gaby and Buckley, 2012) were used with the NZYtaq 2× Green Master Mix (NZYTech, Portugal) in 25 µL reactions. Conditions for PCR amplification were as follows: initial denaturing at 95 °C for 10 min, followed by 35 cycles of denaturing at 95 °C for 45 s, annealing at 56 °C for 30 s, and extending at 72 °C for 1 min. A final extension was performed at 72



°C for 10 min. *Rahnella aquatilis* M72troncoD (HQ538817; Proença et al., 2010) was used as a positive control, and *Escherichia coli* ATCC 25922 was used as a negative control (Gaby and Buckley, 2012). The resulting PCR products were observed in a 2 % agarose gel electrophoresis.

## Results

### Metal(loid) quantification of sediments and plant tissue

Five metals (Cr, Ni, Cu, Zn, and Hg) and the metalloid As were quantified from samples collected at sites C, E, and B, in sediment samples from monospecific stands of *H. portulacoides*, in AG and BG tissues. The results are presented in Table 3.1. For some replicates of Hg concentration determination in sites C and E, the obtained values were below detection limits ( $< 2 \mu\text{g L}^{-1}$ ). In these cases, half of the lowest detectable concentration was used in statistical analyses.

**Table 3.1** Metal(loid) concentrations in sediment and plant tissue samples. Concentrations are given as  $\text{mg kg}^{-1}$  (fresh weight) and presented as mean of three replicates  $\pm$  standard deviation.

Sample	Site	Cr	Ni	Cu	Zn	As	Hg
Sediment							
	C	4.97 $\pm$ 0.67	3.01 $\pm$ 0.39	3.44 $\pm$ 0.30	12.35 $\pm$ 1.51	3.05 $\pm$ 0.37	0.02 $\pm$ 0.01
	E	14.16 $\pm$ 2.05	8.96 $\pm$ 1.24	11.47 $\pm$ 1.60	53.64 $\pm$ 8.78	13.72 $\pm$ 2.07	0.57 $\pm$ 0.01
	B	13.74 $\pm$ 1.90	9.13 $\pm$ 1.32	15.43 $\pm$ 2.10	67.82 $\pm$ 10.75	20.93 $\pm$ 3.34	2.48 $\pm$ 0.17
BG tissues							
	C	0.76 $\pm$ 0.31	0.51 $\pm$ 0.22	1.82 $\pm$ 0.65	8.69 $\pm$ 3.82	0.86 $\pm$ 0.55	0.01 $\pm$ 0.00
	E	1.19 $\pm$ 0.32	0.96 $\pm$ 0.15	2.54 $\pm$ 0.89	16.08 $\pm$ 7.66	8.45 $\pm$ 7.94	0.04 $\pm$ 0.01
	B	1.89 $\pm$ 0.20	1.60 $\pm$ 0.28	3.24 $\pm$ 0.38	22.47 $\pm$ 3.92	4.20 $\pm$ 0.80	0.25 $\pm$ 0.03
AG tissues							
	C	0.69 $\pm$ 0.21	0.47 $\pm$ 0.23	1.77 $\pm$ 1.57	5.31 $\pm$ 1.59	0.40 $\pm$ 0.13	0.01 $\pm$ 0.00
	E	0.53 $\pm$ 0.24	0.21 $\pm$ 0.06	0.33 $\pm$ 0.10	5.52 $\pm$ 1.72	0.35 $\pm$ 0.02	0.01 $\pm$ 0.00
	B	0.82 $\pm$ 0.15	0.52 $\pm$ 0.15	0.53 $\pm$ 0.12	7.26 $\pm$ 2.19	1.03 $\pm$ 0.25	0.07 $\pm$ 0.02

AG, aboveground; BG, belowground.

Sediment samples collected from sites E and B presented concentrations at least 3-fold higher than samples from site C for all metal(loid)s tested ( $p$ -value  $< 0.05$ ). For As and Hg, a contamination gradient could be observed, where site B (closest to the effluent discharge) had the most contaminated sediment samples and site E presented intermediate contamination ( $p$ -value  $< 0.05$ ). Remarkably, sediment Hg concentration was, on average, 31 (site E) to 135 (site B) times higher than in samples from site C. All sediment samples were more contaminated than plant tissue samples, in all sites, for all tested metal(loid)s ( $p$ -value  $< 0.05$ ), except Hg in site C, where the differences in the metal concentration did not present statistical significance.

Considering metal(loid) contents in each site separately, we observed that in site B, metal(loid)s concentration in AG tissues were significantly lower than that in BG tissues. The same was observed for Ni in site E ( $p$ -value < 0.05). Remarkably, Hg contamination of BG tissues in site B was 33-fold higher than in site C.

Considering metal(loid) contents in each type of tissue separately, we saw that BG tissues were significantly more contaminated with Cr, Ni, and Hg in site B than in site C. In the specific cases of As in AG tissues and Hg in both tissues, a significant gradient was observed: C < E < B ( $p$ -value < 0.05).

### **Phylogenetic diversity of endophytic bacteria, community structure, and multivariate analyses**

Endophytic bacteria from sites C, E, and B were isolated from surface sterilized AG and BG tissues from eleven plants, in TSA, MA, and R2A culture media. Extensive results for CFU g<sup>-1</sup> (FW) for all sampled plants are stated in Supplementary Table S3.1. AG tissues from site C presented significantly more CFU g<sup>-1</sup> than BG tissues ( $p$ -value < 0.05). The opposite was observed in sites E ( $p$ -value < 0.05) and B (only for TSA culture medium;  $p$ -value < 0.05).

The selection of morphologically distinct colonies from all culture media, plant tissues, and sampling sites resulted in a collection of 665 endophytic isolates. Molecular typing of the collection yielded 467 representative isolates (150 from CAG community, 74 from CBG, 18 from EAG, 94 from EBG, 33 from BAG, and 97 from BBG), for which partial sequences of 16S rRNA gene were obtained. Five isolates did not yield quality sequences after several attempts and were excluded from the analysis, resulting in 462 amplification products. Taxonomic classification yielded matches with similarity ranging from 92.79 to 100 % similarity. Twenty-nine of 462 isolates (6.3 %) exhibited an identity percentage lower than 97 with the closest match, suggesting the potential for novel species in the collection. Taxonomic information was then affiliated to the respective 658 isolates.

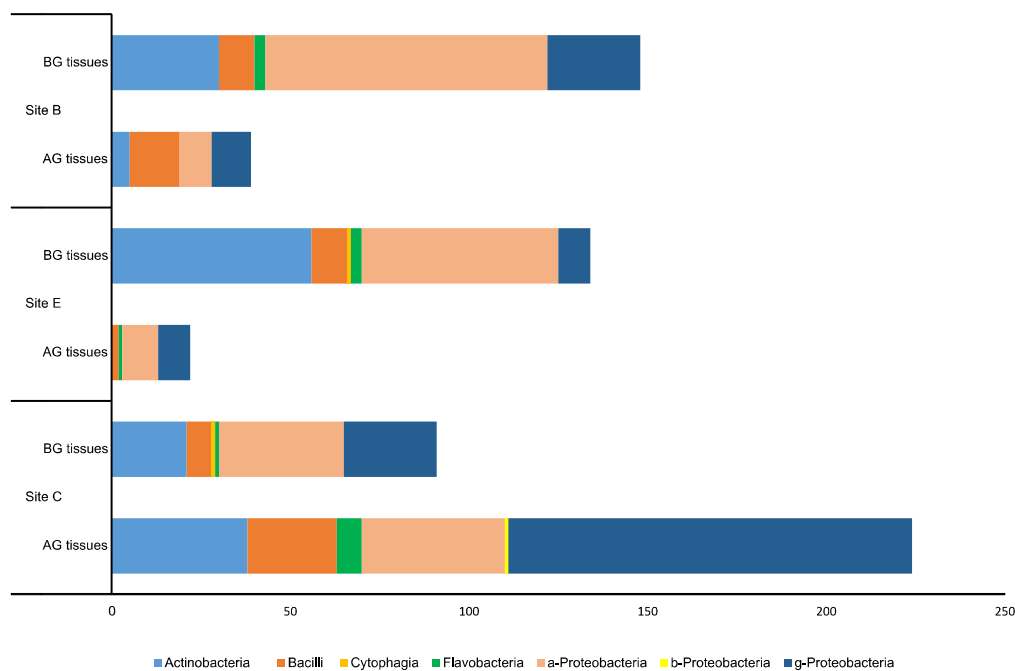
The 658 isolates were distributed amongst class-level identification and grouped according to isolation site and tissue in Figure 3.1. The phylum Proteobacteria (mainly classes Alphaproteobacteria and Gammaproteobacteria) clearly dominated the collection of endophytic bacteria (64.29 % of 658 isolates; Figure 3.1 and Supplementary Table S3.2), followed by Actinobacteria (22.8%) and Firmicutes (10.33%). The phylum Bacteroidetes was also present, but

less represented (2.58 %). Figure 3.1 also shows that Alphaproteobacteria are highly represented in all six communities, and that Gammaproteobacteria, although present in all communities, are most abundant in site C. Isolates from classes Bacilli and Actinobacteria are present in all communities, except Actinobacteria in the community with the lower number of isolates (EAG, 22 isolates).

The most common families of bacteria in the collection were *Vibrionaceae* (51 of 658 isolates, 7.75 %) *Microbacteriaceae* (85/658 isolates, 12.92 %), *Erythrobacteraceae* (87/658 isolates, 13.22 %), and *Halomonadaceae* (98/658 isolates, 14.89 %). The complete distribution of the 658 isolates across 79 genera in all six communities is presented in Supplementary Table S3.2. Overall, the most abundant genera were *Altererythrobacter*, *Marinilactibacillus*, *Microbacterium*, *Salinicola*, and *Vibrio*, with 49, 42, 67, 58, and 51 isolates, respectively. *Altererythrobacter* was the only genus present in all six communities. Isolates belonging to the genera *Altererythrobacter*, *Bacillus*, *Citromicrobium*, *Demequina*, *Labrenzia*, *Marinilactibacillus*, *Martellella*, *Microbacterium*, *Micrococcus*, *Pseudomonas*, *Sphingorabdus*, *Stakelama*, *Vibrio*, and *Zunongwangia* were present in all sites. Isolates belonging to 29 of the 79 genera were present in both BG and AG tissues.

Notably, all 58 isolates (8.8 % of the collection) associated with the genus *Salinicola* were exclusively isolated from site C. *Hoeflea*, *Labrenzia*, and *Microbacterium* were exclusive and considerably dominant in BG tissues, with 17 (2.6 % of the collection), 22 (3.3 %), and 67 (10.2 %) isolates, respectively.

An OTU table was built with abundances across all samples, using the RDP pyrosequencing pipeline. The OTUs were defined considering a distance limit of 0.03 (97 % similarity) and their distribution across communities is stated in Table 3.2. Analysis of indicator OTUs showed that the OTUs, here annotated with the corresponding genera, that contributed most for the characterization of each sampling site were as follows: OTU\_005 (*Salinicola* spp., 58 isolates,  $p$ -value < 0.05) for site C, and OTU\_077 (*Bacillus* spp., 11 isolates,  $p$ -value < 0.05) and OTU\_028 (*Rhizobium* spp., 6 isolates,  $p$ -value < 0.05) for site B. This analysis also showed that indicator OTUs for the isolation tissue were as follows: OTU\_104 (*Zunongwangia* spp., 5 isolates,  $p$ -value < 0.05) for AG tissues, and OTU\_033 (*Labrenzia* spp., 22 isolates,  $p$ -value < 0.05) and OTU\_219 (*Microbacterium* spp., 44 isolates,  $p$ -value < 0.05) for BG tissues.



**Figure 3.1** Distribution of 658 endophytic isolates by class amongst isolation sites (B, E and C) and plant tissues. AG, aboveground; BG, belowground; a-, b- and g-Proteobacteria refers to Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria, respectively.

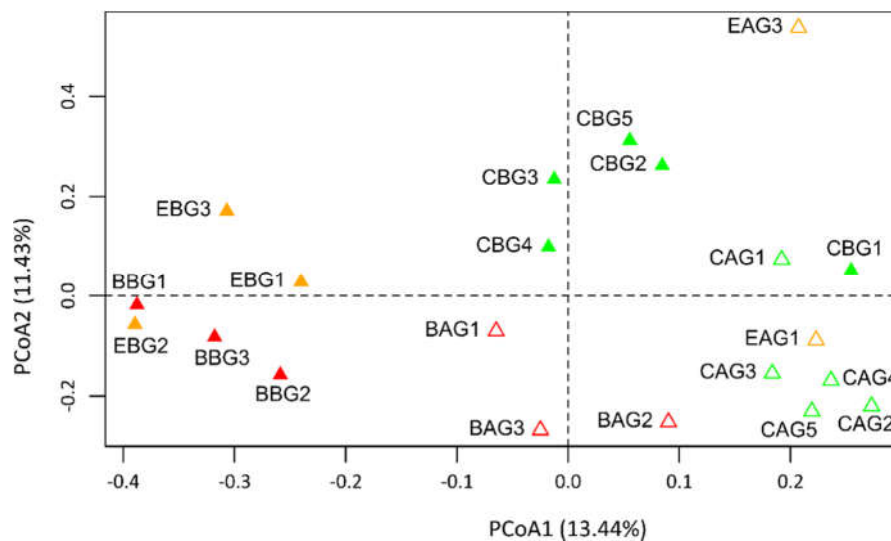
A Bray-Curtis dissimilarity matrix was computed on a square root-transformed OTU table which allowed comparison of bacterial communities of all samples. The resulting matrix was tested for significant differences with PERMANOVA (*adonis* function in R). The interaction between isolation site and tissue was significant (*adonis*:  $F_{2,15} = 1.38$ ,  $R^2 = 0.12$ ,  $p$ -value < 0.05), as was the test for each independent variable: isolation site (*adonis*:  $F_{2,18} = 1.65$ ,  $R^2 = 0.14$ ,  $p$ -value < 0.01) and tissue (*adonis*:  $F_{1,19} = 2.08$ ,  $R^2 = 0.09$ ,  $p$ -value < 0.01). However, only 23 % of the variation in the bacterial composition was explained by both variables and 12 % by their interaction.

A PCoA plot was generated based on the Bray-Curtis dissimilarity matrix to compare all bacterial communities (Figure 3.2). The PCoA plot shows a clear gradient across axis PCoA1 which coincides with the presence of contamination, as samples with most contamination (BG tissues from sites B and E, see Table 3.1) are grouped in the left side of the plot. Analysis of sample scores across PCoA1 revealed that there are significant differences in OTU composition in isolation sites (Kruskal-Wallis chi-squared = 7.66,  $df = 2$ ,  $p$ -value < 0.05) and in isolation tissues (Wilcoxon rank sum test:  $W = 93$ ,  $p$ -value < 0.01). This was not observed in PCoA2 for either isolation sites (ANOVA:  $F_{2,18} = 2.53$ ,  $p$ -value > 0.1) nor tissues (*t* test:  $t = -1.8$ ,  $df = 14.97$ ,  $p$ -value > 0.05).

**Table 3.2** Analysis of 16S rRNA gene partial sequences with RDP pipeline. Sequence and operational taxonomical units (OTUs) information for all isolation sites (C, E and B) and tissues are listed. Clustering was based on a 3 % divergence cutoff. Average values for Shannon's  $H'$  diversity index and Pielou's evenness, computed with R, are also presented for each community.

	CAG	CBG	EAG	EBG	BAG	BBG
Sequence information						
Number of isolates	224	91	22	134	39	148
Average 16S rRNA gene sequence length (bp)	851	876	961	837	937	896
Number of OTUs	52	31	10	41	18	38
Indices						
Average Shannon's $H'$	2.08	2.18	1.20	2.66	1.64	2.57
Average Pielou's Evenness	0.84	0.90	0.76	0.89	0.92	0.88

AG, aboveground; BG, belowground.



**Figure 3.2** Principal coordinates analysis (PCoA) plot based on 16S rRNA gene partial sequences from twenty one endophytic bacterial samples from all isolation sites and tissues. Distances between samples were calculated with Bray-Curtis dissimilarity measure on a square root-transformed OTU table. PCoA axes PCoA1 and PCoA2 are plotted with the corresponding percentages of explained variation. CAG, aboveground (AG) tissues from site C; CBG, belowground (BG) tissues from site C; EAG, AG tissues from site E; EBG, BG tissues from site E; BAG, AG tissues from site B; BBG, BG tissues from site B.

### Enzymatic activity and plant growth promotion traits in endophytic bacteria

Assays for enzymatic activities and PGP traits were successful for 89–100 % of the representative isolates. Test success for siderophore production was 85.2 %; the remaining isolates ( $n = 69$ ) were not able to grow in TSA medium, only MA, where a reaction occurred with the CAS-overlay and interfered with the interpretation of the result. Screening results are stated in Table 3.3 and the extensive list of results for all isolates is in Supplementary Table S3.3.

Only 5.6 to 22.1 % of isolates tested positive for pectinolytic activity. A great percentage of isolates tested positive for all other enzymatic assays (up to 73.2 %), especially IAA production and ACC deaminase activity (98.9 and 51.2 %, respectively). IAA production was detected in 464 isolates, with concentrations ranging from 0.48 to 206  $\mu\text{g mL}^{-1}$ . Seventeen IAA producers presented high levels of IAA concentration, over 100  $\mu\text{g mL}^{-1}$ , 15 of which belonged to Proteobacteria phylum (5 *Pseudomonas* sp.).

**Table 3.3** Screening of bacterial endophytes for enzymatic activity and plant growth promotion traits. Number and percentage of isolates that were successfully tested out of 467 representative isolates (“Tested”), isolates that tested positive (“Positive”) and isolates that tested negative (“Negative”) in traits assays.

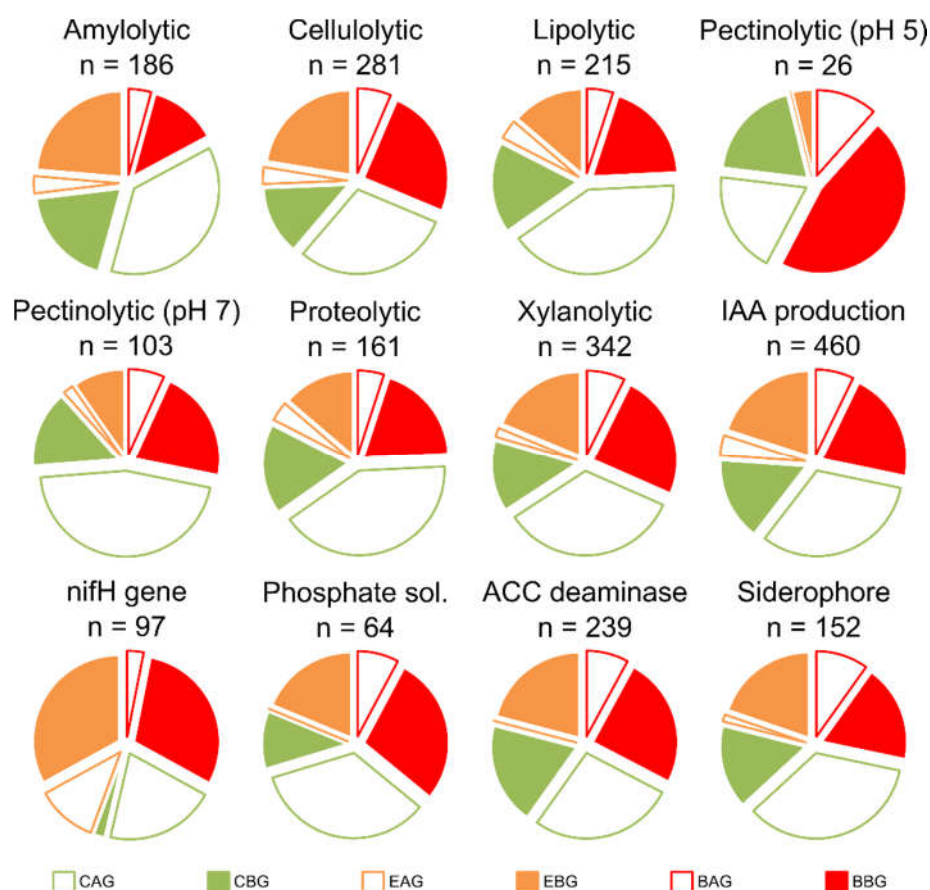
	Enzymatic activity						
	Amylolytic	Cellulolytic	Lipolytic	Pectinolytic (pH 5.0)	Pectinolytic (pH 7.0)	Proteolytic	Xylanolytic
Tested	450	442	416	72	230	446	450
Positive	186 (39.8%)	281 (60.2%)	215 (46.0%)	26 (5.6%)	103 (22.1%)	161 (34.5%)	342 (73.2%)
Negative	264 (56.5%)	161 (34.5%)	201 (43.0%)	46 (9.9%)	127 (27.2%)	285 (61.0%)	108 (23.1%)
	PGP trait						
	ACC deaminase activity	IAA production	Presence of <i>nifH</i> gene	Phosphate solubilization	Siderophore production		
Tested	465	465	467	465	398		
Positive	239 (51.2%)	460 (98.5%)	97 (20.8%)	64 (13.7%)	152 (32.5%)		
Negative	226 (48.4%)	5 (1.1%)	370 (79.2%)	401 (85.9%)	246 (52.7%)		

PGP, plant growth promotion.

Isolates belonging to six genera (one *Chromohalobacter*, one *Halomonas*, one *Marinomonas*, one *Marinospirillum*, one *Micrococcus*, and seven *Vibrio*) tested positive for amylolytic, cellulolytic, lipolytic, pectinolytic (at pH 7), proteolytic, and xylanolytic activities. From the 97 isolates that tested positive for *nifH* amplification, 74 belonged to the Proteobacteria phylum, including 19 *Vibrio* sp. isolates. The most noteworthy genera were *Altererythrobacter* (from the 31 isolates of this genus, 29 were lipolytic and 24 xylanolytic, 22 presented both of these activities), *Marinilactibacillus* (24 cellulolytic, 18 proteolytic and 39 xylanolytic isolates, out of 41 isolates from this genus), *Microbacterium* (32 amylolytic, 30 cellulolytic and 29 xylanolytic isolates, from 39 *Microbacterium* sp. isolates), and *Vibrio* (14 isolates producing every enzyme tested, from a total of 33 *Vibrio* sp. isolates). Isolates associated to five genera (one *Altererythrobacter*, one *Ensifer*, one *Oceanibulbus*, one *Stakelama*, and two *Vibrio*) displayed positive results for all tested PGP traits assays (Supplementary Table S3.3).

The community-based and class-level distribution of representative isolates that tested positive for the assayed enzymatic activities and PGP traits are represented in Figure 3.3 and Supplementary Figure S3.1, respectively. The class-level distribution of isolates that tested positive in all assayed enzymatic activities and PGPs was a reflection of the class-level and community-based distribution of all isolates of our collection (Figure 3.1 and Table 3.2, respectively), i.e., the classes with higher representability in the endophytic collection were also those that presented more isolates that tested positive for the assayed enzymes and traits. This suggests that the probability of the presence of traits in a given class depends on the number of isolates in that class, instead of depending on the site or tissue of isolation.





**Figure 3.3** Distribution of isolates that tested positive for enzymatic activity assays and plant growth promotion traits from 467 representative isolates. Isolates are grouped according to isolation community (site and tissue). *n* represents the number of isolates that tested positive for each trait. Amylolytic, cellulolytic, lipolytic, pectinolytic (at pH 5.0 and 7.0), proteolytic, and xylanolytic activities are represented; ACC deaminase (as growth in DF + ACC medium), IAA production, presence of *nifH* gene, ability to solubilize phosphate, and siderophore production are also represented. CAG, aboveground (AG) tissues from site C; CBG, belowground (BG) tissues from site C; EAG, AG tissues from site E; EBG, BG tissues from site E; BAG, AG tissues from site B; BBG, BG tissues from site B.

## Discussion

The main aim of this study was to elucidate the structure, diversity, and function of culturable endophytic bacterial communities from the salt marsh plant *H. portulacoides*. To our knowledge, this is the first study regarding the isolation, identification, and functional characterization of endophytic bacteria in healthy specimens of the halophyte *H. portulacoides*. Bacteria were isolated from AG and BG tissues from the halophyte, collected from two sites contaminated with industrial effluents and one non-contaminated control site. Several authors have verified that a gradient of contamination exists, from Laranjo Bay to nearshore, especially regarding sediment Hg contamination (reviewed in Pereira et al., 2009). In the present study, sediment samples presented higher metal(loid) contents in sites closest to the industrial effluent discharges (sites B and E). In fact, for As and Hg, a significant concentration gradient could be observed, where content in B > E > C. These metal(loid)s were among the known main contaminants in the industrial effluents in Laranjo Bay, as is stated in the report by Costa and Jesus-Rydin (2001).

Several studies have found that *H. portulacoides* mainly accumulates metals in the BG tissues and that BG to AG translocation is very low (e.g., Valega et al., 2008a; Valega et al., 2008b; Anjum et al., 2011). The same pattern was observed in BG and AG tissues in this study, for all tested metals in site B (up to 6 times more metal content than in AG tissues) and for Ni in site E (average of 4.5 times more Ni in BG than in AG tissues). This allocation pattern was described by Lozano-Rodríguez et al. (1997) as a strategy for protecting the more sensitive AG tissues from the effects of metal-induced stress.

We analyzed a much higher number of isolates than what is usually observed in studies focusing endophytic bacteria. This could be attributed to the fact that we used three culture media with different nutritional compositions and that the plate incubation period for the isolation of the endophytic bacteria was long (72 h). Sample sizes were considerably smaller for EAG and BAG communities and, as a consequence of this unbalance dataset, analyzing indicator OTUs may raise concerns. However, when using the parameter “IndVal.g” in the *indicspecies* package in R, there is a correction for unequal group sizes in the calculations for indicator species (OTUs in our case). The quantity “A,” used in calculating the IndVal, is obtained by using the mean abundance of the species in the target site group divided by the sum of the mean abundance values over all groups. In using the sum of the mean abundance values over all groups, instead of using real abundances from all groups, the unequal sizes of the site groups are controlled for (De Caceres and Legendre, 2009). Regarding the statistical analysis performed based on a dissimilarity measure (Bray-Curtis), we

minimized the most relevant intrinsic bias of this measure (giving too much importance to dominant OTUs and overlooking rare ones) by applying a square root transformation to our OTU abundance table prior to obtaining the Bray-Curtis dissimilarity matrix. The selection of the index at hand was based on (i) its applicability to assess differences between communities using abundance data, (ii) its wide use in the field of ecology and microbiology for this purpose, and (iii) its suitability to answer our questions.

We found considerable culturable bacterial diversity in the internal tissues of *H. portulacoides* from all sampled sites. It is widely accepted that the culturable fraction of plant-associated microbes corresponds to a small fraction of the real diversity, as a consequence of growth medium inadequacy (Tanaka et al., 2014), and limitations to reproduce optimal growth conditions. Nevertheless, the culture-dependent approach allows for extensive functional characterization of the isolates in the collection, which is not possible with culture-independent approaches. Additionally, traditional methods of culture-dependent techniques are useful for the determination of the overall effect of metal contaminants in bacterial communities (Ellis et al., 2001).

As stated above, in our study, the number of isolates obtained from EAG and BAG samples was considerably lower than for other samples. Since the method for obtaining endophytic bacteria includes a harsh surface sterilization process, the viability of endophytes in the samples could be compromised. However, we do not believe this to be the main reason for such a disparity in the number of isolates in the different samples, since the same process was applied to all samples. The isolation efforts were the same for all samples, and as such, the difference in numbers of obtained endophytic isolates may be a reflection of these AG tissues being less populated, since the highest bacterial densities are usually observed in BG tissues, and decrease toward AG tissues (Lodewyckx et al., 2002). In fact, we observed a significantly lower average number of CFU per gram (FW) obtained from EAG and BAG samples, when compared with the respective BG tissues (Supplementary Table S3.1). This pattern was not, however, observed for AG and BG tissues from site C, where the CAG community presented a significantly higher average number of CFU per gram (FW) than CBG. This suggests that the degree of soil and plant tissue contamination could also play a role in the average number of culturable bacteria. This latter effect has been previously reported for bacterial soil communities, where the presence of metal contaminants negatively influenced the ability of readily culturable bacteria to replicate in laboratory media (Ellis et al., 2003).

Proteobacteria, Actinobacteria, Firmicutes, and Bacteroidetes were the phyla recognized in the endophytic collection from this study. This result is in accordance with studies describing the

phyllosphere and endosphere assemblages in other plant hosts (reviewed in Bulgarelli et al., 2013), including halophytes (Mora-Ruiz et al., 2015), and the rhizosphere of *H. portulacoides* (Oliveira et al., 2014b). The genera *Altererythrobacter*, *Marinilactibacillus*, *Microbacterium*, *Salinicola*, and *Vibrio* were the most abundant in our collection. To our knowledge, this is the first time an isolate belonging to the genus *Marinilactibacillus* was obtained from the endosphere.

Isolates affiliated to the genus *Altererythrobacter* have been previously identified in the endosphere (e.g., Liu et al., 2014). In our collection, 73.5 % of isolated *Altererythrobacter* spp. were collected from contaminated sites (E and B). Some species of *Altererythrobacter* are described as able to tolerate metals (Wu et al., 2014), a trait that could be useful in the endosphere of plants colonizing contaminated locations.

*Marinilactibacillus* spp. have been detected in marine environments (e.g., Toffin et al., 2005) and detected with culture-independent methods in olive fermentation brine (Lucena-Padrós et al., 2015). Isolates belonging to this genus have also been isolated from many types of cheese, including as part of consortia that were able to reduce the counts of the human pathogen *Listeria monocytogenes* (e.g., Roth et al., 2010), indicating antimicrobial activities which could be useful in the endosphere environment.

Isolates from the genus *Microbacterium*, exclusively isolated from BG tissues in this study, have been thoroughly characterized and detected in many environments, including the rhizosphere of ginseng (Kim et al., 2015) and the endosphere of crops and prairie plants (Zinniel et al., 2002). Recently, three novel species from this genus were described from the endosphere of *H. portulacoides* (Alves et al., 2014; Alves et al., 2015). PGP traits have been previously associated with this genus (see below), suggesting a high importance in the endosphere.

Isolates from the genus *Salinicola* have been previously obtained from diverse environments (e.g., deep sea surface sediments, Romanenko et al., 2013) including the endosphere of a halophyte (Mora-Ruiz et al., 2015). To the best of our knowledge, PGP traits have not been explored in this genus. The prosperity of *Salinicola* spp. in marine environments could be related to their presence in the endosphere of the halophyte at study.

*Vibrio* spp. have been abundantly detected in aquatic environments and organisms (Thompson et al., 2004) and in the endosphere (e.g., seagrass, Jose et al., 2014). PGP traits have been associated with this genus (see below), suggesting that this widely known pathogen in humans and other animals may have a positive role in the endosphere.

The genera *Pseudomonas* and *Micrococcus*, less represented in this study (15 and 9 isolates, respectively), were previously detected in the endosphere of *H. portulacoides* in a study regarding hydrocarbon-degrading bacteria (Oliveira et al., 2014a).

From the 462 identified isolates, 29 presented less than 97 % similarity with the closest 16S rRNA gene sequence in GenBank. These isolates may potentially belong to novel species, revealing the endosphere of *H. portulacoides* as a source for undescribed phylogenetic groups.

As reviewed in Bulgarelli et al. (2013) for many plant hosts and discussed in Oliveira et al. (2014b) for *H. portulacoides*, plant-related factors contribute to the composition of plant-associated bacterial communities. In this work, 14 genera were isolated from all sites, suggesting the presence of a common structure associated to the studied halophyte. On the other hand, differences were observed, where considerably abundant genera were exclusively isolated from one site (e.g., *Salinicola* spp. from site C).

In the present study, *Altererythrobacter* spp., *Marinilactibacillus* spp., *Microbacterium* spp., and *Vibrio* spp. isolates were abundantly detected in all sites, suggesting an ability to tolerate metals, which has been previously described in these genera (Wu et al., 2014; Sheng et al., 2008; Jafari et al., 2015) except, to the best of our knowledge, for the genus *Marinilactibacillus*. OTUs belonging to the genera *Bacillus* and *Rhizobium*, albeit less represented in the endophyte collection (11 and 6 isolates, respectively), were considered indicator OTUs for site B, and metal tolerance has previously been associated with these genera (Ma et al., 2015; Aafi et al., 2015). Isolates belonging to *Salinicola* spp. ( $n=58$ ) were exclusively detected in site C, and the indicator OTU for site C was also attributed to this genus. Information on whether *Salinicola* spp. can tolerate metal contamination is lacking in the literature; however, considering the number of isolates obtained from this genus and the extension of our collection, we suggest that it should be further investigated as to its ability as a bioindicator.

Usually, bacteria access plant tissues through natural entry points in the root tips and lateral roots emergence points, as a consequence of root growth. Once inside the plant, endophytes may colonize the plant by migration from BG to AG tissues. Additional penetration may occur through natural openings on leaves such as stomata, or through leaf wounds (reviewed in Lodewyckx et al., 2002). Here, we found that BG and AG tissues shared 20 bacterial genera, suggesting that migration from BG to AG tissues may occur in *H. portulacoides*. On the other hand, isolates belonging to the genera *Hoeflea* (17), *Labrenzia* (22), and *Microbacterium* (67) were only isolated from BG tissues, and the two latter genera were identified as indicator OTUs for these tissues by our indicator

species analysis. The genera *Hoeflea* and *Labrenzia* have been isolated from root endosphere of a salt marsh halophyte (Bibi et al., 2014) and, to the best of our knowledge, have not been detected in AG tissues of plants, as was the case in our study. On the other hand, *Microbacterium* spp. has been detected in AG tissues of *H. portulacoides* (e.g., Alves et al., 2014). An OTU attributed to the genus *Zunongwangia* was considered indicator OTU for AG tissues in our analysis. Although this genus has been isolated from the rhizosphere of mangrove trees (Rameshkumar et al., 2014) and detected by pyrosequencing in the endosphere (Mora-Ruiz et al., 2015), it had not been previously isolated in the endosphere.

We found that the bacterial composition across our collection is significantly affected by the isolation site and tissue, although these factors only account for 23 % of the total variation observed in our data. Nevertheless, a variation in composition could be observed across the gradient of contamination in our sampling sites, suggesting that the contamination levels do, in fact, affect the composition of the culturable endosphere of *H. portulacoides*. Other variation parameters (e.g., soil water content, organic matter content, fine particle content, redox potential, pH, and conductivity; Valega et al., 2008b) could be explored in order to further understand the reasons for composition variation in the culturable endosphere of this halophyte.

Many isolates from the collection obtained in this study tested positive for several enzymatic activities and PGP traits, exposing the endosphere of the halophyte *H. portulacoides* as a novel source of bacteria and bacterial compounds with potential biotechnological applications. The ability to produce these extracellular enzymes and PGP traits is well distributed across phylogenetic groups, as the most represented classes (Figure 3.1) amply exhibited all tested traits, except Bacilli for phosphate solubilization (Supplementary Figure S3.1). This ability is also well represented, overall, in all isolation sites and tissues (Figure 3.3). The ability to produce industrially important extracellular enzymes was abundantly detected in isolates associated to the genera *Altererythrobacter*, *Marinilactibacillus*, *Microbacterium*, and *Vibrio*. To our knowledge, the production of the tested enzymes had not been previously described in *Altererythrobacter* spp. and *Marinilactibacillus* spp.

To the best of our knowledge, the PGP traits tested in the presented study were not previously evaluated in *Altererythrobacter* spp. isolates. Here, we found that isolates from this genus tested positive for one to six extracellular enzymatic activities and presented one to five PGP traits. Overall, all tested enzymes and traits were produced by isolates belonging to his genus.

Traits for *Marinilactibacillus* spp. isolates from other studies have not been deeply characterized. Here, we found that *Marinilactibacillus* spp. isolates tested positive for one to five extracellular enzymatic activities, and presented one to four PGP traits. Considering all the isolates from this genus, all of the tested enzymes and traits were produced, except pectinase at pH 5.

The ability to promote plant growth has been previously detected in the genus *Microbacterium* with traits as ACC deaminase (Sheng et al., 2008), IAA production, phosphate solubilization (Kukla et al., 2014), and siderophore production (Pereira et al., 2013). The presence of the *nifH* gene was also detected (Zakhia et al., 2006). Production of protease and lipase (Pereira et al., 2013), xylanase, amylase (Park et al., 2006), cellulase (Kukla et al., 2014), and pectinase (Vinod et al., 2014) has been characterized in *Microbacterium* spp.. Antimicrobial activity and resistance to metals have also been reported in this genus (Sheng et al., 2008; Passari et al., 2015). In the present study, one to five of the abovementioned extracellular enzymes were also detected in all *Microbacterium* spp. isolates, as were one to four PGP traits. Considering all the isolates from this genus, all of the tested enzymes and traits were present.

Previously studied *Vibrio* spp. exhibited enzymatic activities such as proteolytic, cellulolytic, amylolytic (Jose et al., 2014), xylanolytic (Kiyohara et al., 2005), and lipolytic (Ray et al., 2012), showed potential for plant growth promotion by solubilizing phosphate, fixing atmospheric nitrogen, producing IAA (Jose et al., 2014), and siderophores (Thompson et al., 2004). *Vibrio* spp. isolates from our collection tested positive for these activities and traits and, in addition, pectinolytic activity (at both pH=5.0 and 7.0) and ACC deaminase activity. Other studies have also detected antimicrobial activity (Kumar and Nair, 2007) and considerable Hg resistance and ability for Hg bioremediation in *Vibrio* spp. isolates (Jafari et al., 2015).

IAA production was detected in 460 isolates, 17 of which produced over 100  $\mu\text{g mL}^{-1}$ . The most represented IAA producing genus is *Pseudomonas* with five isolates, which has previously been associated with IAA producing potential in rice root exudates (Karnwal, 2009). Phosphate solubilization was detected in 64 isolates, 10 of which belong to the genus *Microbacterium*, and 15 belong to the genus *Salinicola*. While *Microbacterium* spp. have been associated with phosphate solubilizing activity (Kukla et al., 2014), this is, to our knowledge, the first report of such activity in the *Salinicola* genus. Using PCR amplification, 20.8 % of representative isolates tested positive for the presence of the *nifH* gene. Of the 97 *nifH*-positive isolates, 63 originated from BG tissues and 17 belonged to the genus *Vibrio*, which has previously been associated with this trait (Jose et al., 2014). According to Penrose and Glick (2003), the ability to use ACC as a sole nitrogen source is a

consequence of the activity of the enzyme ACC deaminase. From the 239 isolates that were able to grow in DF minimal medium supplemented with ACC, 155 (64.9 %) originated from BG tissues, and 60 % of isolates sampled in site B tested positive for this activity. Isolates associated with the genera *Microbacterium* (33) and *Altererythrobacter* (25) contributed the most for this PGP trait. Isolates from *Microbacterium* genus have previously been associated with ACC deaminase activity (Sheng et al., 2008); however, to our knowledge, this is the first description of such activity in the genus *Altererythrobacter*. Siderophore production was observed in 152 isolates (32.5 % of the collection), 19 of which belong to the genus *Vibrio*, which had been previously linked to this activity (Thompson et al., 2004). Nineteen *Marinilactibacillus* spp. isolates also tested positive, which is, to our knowledge, the first time such activity was described in this genus.

This work revealed that the endosphere of *H. portulacoides* represents a bacterial hotspot harboring a diverse microbiota, including putative novel species and compounds with biotechnological relevance. Endophytic bacteria were isolated from AG and BG tissues of the halophyte from contaminated and non-contaminated sites. Seventy-nine different genera were found using traditional cultivation methods. The phylogenetically diverse collection presented structural differences among sampling tissues and sites, and many of the isolates revealed their ability to produce extracellular enzymes and PGP traits. The PGP potential of these endophytes should be further characterized, aiming toward enhancing phytoextraction of metal-contaminated soils by *H. portulacoides*. The possibility of monitoring *Salinicola* spp. as a bioindicator of lack of contamination should also be further studied. Metal resistance and antimicrobial properties of these isolates should also be assessed to further characterize functional capabilities of the collection.

## Acknowledgments

This work was financed by the European Funds through COMPETE and by National Funds through the Portuguese Foundation for Science and Technology (FCT) within project PhytoMarsh (PTDC/AAC-AMB/118873/2010–FCOMP-01-0124-FEDER-019328). The authors acknowledge FCT financing to CESAM (UID/AMB/50017/2013) and Institute for Research in Biomedicine (iBiMED–UID/BIM/04501/2013), Artur Alves (FCT Investigator Programme–IF/00835/2013), Isabel Henriques (FCT Investigator Programme–IF/00492/2013), and Cátia Fidalgo (PhD grant–SFRH/BD/85423/2012). Authors acknowledge Paula Castro and Diogo Proença for kindly providing positive and negative control strains used in this study.



## References

- Aafi, N. E., Saidi, N., Maltouf, A. F., Perez-Palacios, P., Dary, M., Brhada. F. & Pajuelo, E. (2015). Prospecting metal-tolerant rhizobia for phytoremediation of mining soils from Morocco using *Anthyllis vulneraria* L. *Environ Sci Pollut Res* **22**, 4500–4512.
- Altschul, S.F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Alongi, D. M. (1998). Mangroves and salt marshes. In: Kennish MJ, Lutz PL (eds) Coastal ecosystem processes. CRC, Florida, 419 pp. ISBN 0-8493-8426-5
- Alves, A., Correia, A., Igual, J. M. & Trujillo, M. E. (2014). *Microbacterium endophyticum* sp. nov. and *Microbacterium halimionae* sp. nov., endophytes isolated from the salt-marsh plant *Halimione portulacoides* and emended description of the genus *Microbacterium*. *Syst Appl Microbiol* **37**, 474–479.
- Alves, A., Riesco, R., Correia, A. & Trujillo, M. E. (2015). *Microbacterium proteolyticum* sp. nov. isolated from roots of *Halimione portulacoides*. *Int J Syst Evol Microbiol* **65**, 1794–1798.
- Ando, S., Goto, M., Meunchang, S., Thongra-ar, P., Fujiwara, T., Hayashi, H. & Yoneyama, T. (2005). Detection of *nifH* sequences in sugarcane (*Saccharum officinarum* L.) and pineapple (*Ananas comosus* [L.] Merr.). *Soil Sci Plant Nutr* **51**, 303–308.
- Anjum, N. A., Ahmad, I., Válega, M., Pacheco, M., Figueira, E., Duarte, A. C. & Pereira, E. (2011). Impact of seasonal fluctuations on the sediment mercury, its accumulation and partitioning in *Halimione portulacoides* and *Juncus maritimus* collected from Ria de Aveiro coastal lagoon (Portugal). *Water Air Soil Pollut* **222**, 1–15.
- Berg, G., Grube, M., Schlöter, M. & Smalla, K. (2014). Unraveling the plant microbiome: looking back and future perspectives. *Front Microbiol* **5**, 148.
- Bibi, F., Jeong, J. H., Chung, E. J., Jeon, C. O. & Chung, Y. R. (2014). *Labrenzia suaedae* sp. nov., a marine bacterium isolated from a halophyte, and emended description of the genus *Labrenzia*. *Int J Syst Evol Microbiol* **64**, 1116–1122.
- Bouchard, V., Creach, V., Lefeuvre, J. C., Bertru, G. & Mariotti, A. (1998). Fate of plant detritus in a European salt marsh dominated by *Atriplex portulacoides* (L.) Aellen. *Hydrobiologia* **373**, 75–87.
- Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L. & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* **64**, 807–838.
- Carvalho, P. N., Basto, M. C. P., Silva, M. F. G. M., Machado, A., Bordalo, A. A. & Vasconcelos, T. S. D. (2010). Ability of salt marsh plants for TBT remediation in sediments. *Environ Sci Pollut Res* **17**, 1279–1286.

- Costa, C. & Jesus-Rydin, C. (2001).** Site investigation on heavy metals contaminated ground in Estarreja – Portugal. *Eng Geol* **60**, 39–47.
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R. & Tiedje, J. M. (2014).** Ribosomal Database Project: data and tools for high-throughput rRNA analysis. *Nucleic Acids Res* **41**, D633–D642.
- Couto, M. N. P. F. S., Basto, M. C. P. & Vasconcelos, M. T. S. D. (2011).** Suitability of different salt marsh plants for petroleum hydrocarbons remediation. *Chemosphere* **84**, 1052–1057.
- Cox, C. D. (1994).** Deferration of laboratory media and assays for ferric and ferrous ions. *Method Enzymol* **235**, 315–329.
- De Caceres, M. & Legendre, P. (2009).** Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**, 3566–3574.
- Dworkin, M. & Foster, J. W. (1958).** Experiments with some microorganisms which utilize ethane and hydrogen. *J Bacteriol* **75**, 592–603.
- Ellis, R. J., Neish, B., Trett, M. W., Best, J. G., Weightman, A. J., Morgan, P. & Fry, J. C. (2001).** Comparison of microbial and meiofaunal community analyses for determining impact of heavy metal contamination. *J Microbiol Meth* **45**, 171–185.
- Ellis, R. J., Morgan, P., Weightman, A. J. & Fry, J. C. (2003).** Cultivation-dependent and -independent approaches for determining bacterial diversity in heavy-metal-contaminated soil. *Appl Environ Microb* **69**, 3223–3230.
- Figueira, E. & Freitas, R. (2013).** Consumption of *Ruditapes philippinarum* and *Ruditapes decussatus*: comparison of element accumulation and health risk. *Environ Sci Pollut R* **20**, 5682–5691.
- Gaby, J. C. & Buckley, D. H. (2012).** A comprehensive evaluation of PCR primers to amplify the *nifH* gene of nitrogenase. *PLoS ONE* **7**, 7.
- Gordon, S. A. & Weber, R. P. (1951).** Colorimetric estimation of indoleacetic acid. *Plant Physiol* **26**, 192–195.
- Hardoim, P. R., van Overbeek, L. S. & van Elsas, J. D. (2008).** Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* **16**, 463–471.
- Jafari, S. A., Cheraghi, S., Mirbakhsh, M., Mirza, R. & Maryamabadi, A. (2015).** Employing response surface methodology for optimization of mercury bioremediation by *Vibrio parahaemolyticus* PG02 in Coastal Sediments of Bushehr, Iran. *CLEAN* **43**, 118–126.

- Jose, P. A., Sundari, I. S., Sivakala, K. K. & Jebahumar, S. R. D. (2014).** Molecular phylogeny and plant growth promoting traits of endophytic bacteria isolated from roots of seagrass *Cymodocea serrulata*. *Indian J GeoMar Sci* **43**, 571–579.
- Karnwal, A. (2009).** Production of indole acetic acid by fluorescent *Pseudomonas* in the presence of L-tryptophan and rice root exudates. *J of Plant Pathol* **91**, 61–63.
- Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H., Yi, H., Won, S. & Chun, J. (2012).** Introducing EzTaxon: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Micr* **62**, 716–721.
- Kim, Y.-J., Nguyen, N.-L., Hoang, V.-A., Min, J.-W., Hwang, K.-H. & Yang, D.-C. (2015).** *Microbacterium panaciterrae* sp. nov., isolated from the rhizosphere of ginseng. *Int J Syst Evol Micr* **65**, 927–933.
- Kiyohara, M., Sakaguchi, K., Yamaguchi, K., Akari, T., Nakamura, T. & Ito, M. (2005).** Molecular cloning and characterization of a novel  $\beta$ -1,3-xylanase possessing two putative carbohydrate-binding modules from a marine bacterium *Vibrio* sp. strain AX-4. *Biochem J* **388**, 949–957.
- Kukla, M., Płociniczak, T. & Piotrowska-Seget, Z. (2014).** Diversity of endophytic bacteria in *Lolium perenne* and their potential to degrade petroleum hydrocarbons and promote plant growth. *Chemosphere* **117**, 40–46.
- Kumar, N. R. & Nair, S. (2007).** *Vibrio rhizosphaerae* sp. nov., a red-pigmented bacterium that antagonizes phytopathogenic bacteria. *Int J Syst Evol Micr* **57**, 2241–2246.
- Lane, D. J. (1991).** 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds) *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons, New York, pp 115–175.
- Liu, X. L., Liu, S. L., Liu, M., Kong, B. H., Liu, L. & Li, Y. H. (2014).** A primary assessment of the endophytic bacterial community in a xerophilous moss (*Grimmia montana*) using molecular method and cultivated isolates. *Braz J Microbiol* **45**, 163–173.
- Lodewyckx, C., Vangronsveld, J., Porteous, F., Moore, E. R. B., Taghavi, S., Mezgeay, M. & van der Lelie, D. (2002).** Endophytic bacteria and their potential applications. *Crit Rev Plant Sci* **21**, 583–606.
- Lozano-Rodríguez, E., Hernández, L. E., Bonay, P. & Carpena-Ruiz, R. O. (1997).** Distribution of cadmium in shoot and root tissues of maize and pea plants: physiological disturbances. *J Exp Bot* **48**, 123–128.
- Lucena-Padrós, H., Jiménez, E., Maldonado-Barragán, A., Rodríguez, J. M. & Ruiz-Barba, J. L. (2015).** PCR-DGGE assessment of the bacterial diversity in Spanish-style green table-olive fermentations. *Int J Food Microbiol* **205**, 47–53.

- Ma, Y., Oliveira, R. S., Nai, F., Rajkumar, M., Luo, Y., Rocha, I. & Freitas, H. (2015).** The hyperaccumulator *Sedum plumbizincicola* harbors metal-resistant endophytic bacteria that improve its phytoextraction capacity in multi-metal contaminated soil. *J Environ Manage* **156**, 62–69.
- Mora-Ruiz, M. R., Font-Verdera, F., Díaz-Gil, C., Urdiain, M., Rodríguez-Valdecantos, G., González, B., Orfila, A. & Rosselló-Móra, R. (2015).** Moderate halophilic bacteria colonizing the phylloplane of halophytes of the subfamily Salicornioideae (*Amaranthaceae*). *Syst Appl Microbiol* **38**, 406–416.
- Nautiyal, C. S. (1999).** An efficient microbiological growth medium for screening phosphate solubilizing microorganisms. *FEMS Microbiol Lett* **170**, 265–270.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O’Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H. & Wagner, H. (2015).** vegan: Community Ecology Package. R package version 2.2-1. <http://CRAN.R-project.org/package=vegan>
- Oliveira, V., Gomes, N. C. M., Almeida, A., Silva, A. M. S., Simões, M. Q. M., Smalla, K. & Cunha, A. (2014a).** Hydrocarbon contamination and plant species determine the phylogenetic and functional diversity of endophytic degrading bacteria. *Mol Ecol* **23**, 1392–1404.
- Oliveira, V., Gomes, N. C. M., Cleary, D. F. R., Almeida, A., Silva, A. M. S., Simões, M. M. Q., Silva, H. & Cunha, A. (2014b).** Halophyte plant colonization as a driver of the composition of bacterial communities in salt marshes chronically exposed to oil hydrocarbons. *FEMS Microbiol Ecol* **90**, 647–662.
- Park, H. Y., Kim, K. K., Jin, L. & Lee, S.-T. (2006).** *Microbacterium paludicola* sp. nov., a novel xylanolytic bacterium isolated from swamp forest. *Int J Syst Evol Micr* **56**, 535–539.
- Passari, A. K., Mishra, V. K., Saikia, R., Gupta, V. K. & Singh, B. P. (2015).** Isolation, abundance and phylogenetic affiliation of endophytic actinomycetes associated with medicinal plants and screening for their in vitro antimicrobial biosynthetic potential. *Front Microbiol* **6**, 273.
- Penrose, D. M. & Glick, B. R. (2003).** Methods for isolating and characterizing ACC deaminase-containing plant growth-promoting rhizobacteria. *Physiol Plantarum* **118**, 10–15.
- Pereira, M. E., Duarte, A. C., Millward, G. E., Vale, C. & Abreu, S. N. (1998).** Tidal export of particulate mercury from the most contaminated area of Aveiro’s Lagoon, Portugal. *Sci Total Environ* **213**, 157–163.
- Pereira, M. E., Lillebø, A. I., Pato, P., Válega, M., Coelho, J. P., Lopes, C. B., Rodrigues, S., Chachada, A., Otero, M., Pardal, M. A. & Duarte, A. C. (2009).** Mercury pollution in Ria de Aveiro (Portugal): a review of the system assessment. *Environ Monit Assess* **155**, 39–49.
- Pereira, S. I. A., Barbosa, L. & Castro, P. M. L. (2013).** Rhizobacteria isolated from a metal-polluted area enhance plant growth in zinc and cadmium-contaminated soil. *Int J Environ Sci Technol* **12**, 2127–2142.

- Pérez-Miranda, S., Cabirol, N., George-Téllez, R., Zamudio-Rivera, L. S. & Fernández, F. J. (2007).** O-CAS, a fast and universal method for siderophore detection. *J Microbiol Meth* **70**, 127–131.
- Pohlert, T. (2005).** PMCMR: Calculate Pairwise Multiple Comparisons of Mean Rank Sums. R package version 1.1. <http://CRAN.R-project.org/package=PMCMR>
- Proença, D. N., Francisco, R., Santos, C. V., Lopes, A., Fonseca, L., Abrantes, I. M. O. & Morais, P. V. (2010).** Diversity of bacteria associated with *Bursaphelenchus xylophilus* and other nematodes isolated from *Pinus pinaster* trees with pine wilt disease. *PLoS One* **5**, 12.
- R Core Team (2014).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Rajkumar, M., Sandhya, S., Prasad, M. N. V. & Freitas, H. (2012).** Perspectives of plant-associated microbes in heavy metal phytoremediation. *Biotechnol Adv* **30**, 1562–1574.
- Rameshkumar, N., Krishnan, R., Lang, E., Matsumura, Y., Sawabe, T. & Sawabe, T. (2014).** *Zunongwangia mangrovi* sp. nov., isolated from mangrove (*Avicennia marina*) rhizosphere, and emended description of the genus *Zunongwangia*. *Int J Syst Evol Microbiol* **64**, 545–550.
- Ray, A. K., Ghosh, K. & Ringø, E. (2012).** Enzyme-producing bacteria isolated from fish gut: a review. *Aquacult Nutr* **18**, 465–492.
- Romanenko, L. A., Tanaka, N., Kalinovskaya, N. I. & Mikhailov, V. V. (2013).** Antimicrobial potential of deep surface sediment associated bacteria from the Sea of Japan. *World J Microb Biot* **29**, 1169–1177.
- Roth, E., Schwenninger, S. M., Hasler, M., Eugster-Meier, E. & Lacroix, C. (2010).** Population dynamics of two antilisterial cheese surface consortia revealed by temporal temperature gradient gel electrophoresis. *BMC Microbiol* **10**, 74.
- Sessitsch, A., Kuffner, M., Kidd, P., Vangronsveld, J., Wenzel, W. W., Fallmann, K. & Puschenreiter, M. (2013).** The role of plant-associated bacteria in the mobilization and phytoextraction of trace elements in contaminated soils. *Soil Biol Biochem* **60**, 182–194.
- Sheng, X. F., Xia, J. J., Jiang, C. Y., He, L. Y. & Qian, M. (2008).** Characterization of heavy metal-resistant endophytic bacteria from rape (*Brassica napus*) roots and their potential in promoting the growth and lead accumulation of rape. *Environ Pollut* **156**, 1164–1170.
- Tanaka, T., Kawasaki, K., Daimon, S., Kitagawa, W., Yamamoto, K., Tamaki, H., Tanaka, M., Nakatsu, C. H. & Kamagata, Y. (2014).** A hidden pitfall in agar media preparation undermines cultivability of microorganisms. *Appl Environ Microbiol* **80**, 7659–7666.
- Thompson, F. L., Iida, T. & Swings, J. (2004).** Biodiversity of *Vibrios*. *Microbiol Mol Biol R* **68**, 403–431.

- Toffin, L., Zink, K., Kato, C., Pignet, P., Bidault, A., Bienvenu, N., Birrien, J.-L. & Prieur, D. (2005).** *Marinilactibacillus piezotolerans* sp. nov., a novel marine lactic acid bacterium isolated from deep sub-seafloor sediment of the Nankai Trough. *Int J Syst Evol Microbiol* **55**, 345–351.
- Válega, M., Lillebø, A. I., Pereira, M. E., Caçador, I., Duarte, A. C. & Pardal, M. A. (2008a).** Mercury in salt marshes ecosystems: *Halimione portulacoides* as biomonitor. *Chemosphere* **73**, 1224–1229.
- Válega, M., Lillebø, A. I., Pereira, M. E., Duarte, A. C. & Pardal, M. A. (2008b).** Long-term effects of Mercury in a salt marsh: hysteresis in the distribution of vegetation following recovery from contamination. *Chemosphere* **71**, 765–772.
- Versalovic, J., Koeuth, T. & Lupski, J. R. (1991).** Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* **19**, 6823–6831.
- Versalovic, J., Schneider, M., de Bruijn, F. J. & Lupski, J. R. (1994).** Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Method Mol Cell Biol* **5**, 25–40.
- Vinod, V., Kumar, A. & Zachariah, T. J. (2014).** Isolation, characterization and identification of pericarp-degrading bacteria for the production of off-odour-free white pepper from fresh berries of *Piper nigrum* L. *J Appl Microbiol* **116**, 890–902.
- Woerner, L. S. & Hackney, C. T. (1997).** Distribution of *Juncus roemerianus* in North Carolina tidal marshes: the importance of physical and biotic variables. *Wetlands* **17**, 284–291.
- Wu, Y. H., Xu, L., Meng, F. X., Zhang, D. S., Wang, C. S., Oren, A. & Xu, X.-W. (2014).** *Altererythrobacter atlanticus* sp. nov., isolated from deep-sea sediment. *Int J Syst Evol Microbiol* **64**, 116–121.
- Zakhia, F., Jeder, H., Willems, A., Gillis, M., Dreyfus, B. & de Lajudie, P. (2006).** Diverse bacteria associated with root nodules of spontaneous legumes in Tunisia and first report for *nifH*-like gene within the genera *Microbacterium* and *Starkeya*. *Microbial Ecol* **51**, 375–393.
- Zinniel, D. K., Lambrecht, P., Harris, N. B., Feng, Z., Kuczmarski, D., Higley, P., Ishimaru, C. A., Arunakumari, A., Barletta, R. G. & Vidaver, A. K. (2002).** Isolation and characterization of endophytic colonizing bacteria from agronomic crops and prairie plants. *Appl Environ Microb* **68**, 2198–2208.

## Supplementary Material

**Supplementary Table S3.1** Colony forming units (CFUs) per gram of vegetal tissue (fresh weight) in samples from all isolation sites (C, E and B) and tissues, per culture media. Values for CFU g<sup>-1</sup> (FW) are given as the average  $\pm$  standard deviation of two replicates in the countable dilution, for all plants collected in each community.

Plant	Culture medium	CFU g <sup>-1</sup>	
		Average	Standard deviation
CAG-1	MA	5.39E+04	1.97E+04
	R2A	0.00E+00	0.00E+00
	TSA	0.00E+00	0.00E+00
CBG-1	MA	7.58E+03	2.53E+03
	R2A	0.00E+00	0.00E+00
	TSA	0.00E+00	0.00E+00
CAG-2	MA	3.44E+04	2.25E+04
	R2A	3.66E+04	3.61E+04
	TSA	4.65E+04	4.19E+04
CBG-2	MA	3.85E+03	1.54E+03
	R2A	0.00E+00	0.00E+00
	TSA	3.85E+02	3.85E+02
CAG-3	MA	5.81E+04	5.80E+04
	R2A	2.45E+05	2.04E+05
	TSA	2.97E+05	1.74E+05
CBG-3	MA	9.76E+04	2.40E+04
	R2A	1.21E+05	9.62E+02
	TSA	2.70E+05	1.46E+05
CAG-4	MA	1.48E+08	2.73E+07
	R2A	1.84E+08	7.37E+06
	TSA	1.32E+08	2.16E+07
CBG-4	MA	7.71E+06	7.20E+05
	R2A	1.53E+07	1.01E+06
	TSA	1.28E+07	1.79E+06
CAG-5	MA	2.22E+06	1.94E+06
	R2A	2.39E+04	5.98E+03
	TSA	1.59E+04	5.98E+03
CBG-5	MA	3.22E+05	9.79E+04
	R2A	6.99E+03	6.99E+03
	TSA	0.00E+00	0.00E+00
EAG-1	MA	2.67E+04	2.13E+04
	R2A	1.78E+02	1.78E+02
	TSA	1.16E+03	8.00E+02
EBG-1	MA	5.69E+05	5.63E+04
	R2A	1.38E+05	5.85E+03
	TSA	2.09E+04	1.13E+03
EAG-2	MA	0.00E+00	0.00E+00
	R2A	0.00E+00	0.00E+00
	TSA	9.98E+01	9.98E+01
EBG-2	MA	2.72E+05	4.13E+04
	R2A	9.18E+04	1.81E+04
	TSA	1.14E+04	8.42E+03

EAG-3	MA	1.19E+03	1.19E+03
	R2A	0.00E+00	0.00E+00
	TSA	0.00E+00	0.00E+00
EBG-3	MA	5.71E+04	2.07E+04
	R2A	5.67E+03	2.46E+02
	TSA	2.78E+04	1.95E+04
BAG-1	MA	4.92E+02	4.92E+02
	R2A	9.84E+01	9.84E+01
	TSA	1.18E+03	1.18E+03
BBG-1	MA	6.61E+05	5.41E+05
	R2A	1.85E+04	5.53E+03
	TSA	1.15E+05	1.13E+05
BAG-2	MA	7.97E+02	1.99E+02
	R2A	8.96E+02	8.96E+02
	TSA	2.99E+02	2.99E+02
BBG-2	MA	3.37E+05	3.28E+04
	R2A	7.96E+04	8.13E+03
	TSA	6.77E+04	1.90E+04
BAG-3	MA	6.62E+03	6.92E+02
	R2A	4.45E+03	4.94E+02
	TSA	1.98E+03	5.93E+02
BBG-3	MA	1.88E+06	8.54E+05
	R2A	4.12E+05	2.00E+04
	TSA	2.96E+05	3.76E+04

CAG, aboveground (AG) tissues from site C; CBG, belowground (BG) tissues from site C; EAG, AG tissues from site E; EBG, BG tissues from site E; BAG, AG tissues from site B; BBG, BG tissues from site B.



**Supplementary Table S3.2** Phylum, class and genus-level distribution of 658 bacterial endophytic isolates in three sites (C, E and B) and two types of tissues.

Taxonomic rank and taxon	Site C		Site E		Site B		Overall
	AG tissues	BG tissues	AG tissues	BG tissues	AG tissues	BG tissues	
Phylum Actinobacteria							
Class Actinobacteria	38 (5.78%)	21 (3.19%)	0	56 (8.51%)	5 (0.76%)	30 (4.56%)	150 (22.80%)
<i>Aeromicrobium</i>	0	0	0	2	0	0	2
<i>Curtobacterium</i>	5	0	0	1	0	0	6
<i>Demequina</i>	3	5	0	6	2	11	27
<i>Dietzia</i>	11	5	0	3	0	0	19
<i>Frigoribacterium</i>	1	0	0	0	0	0	1
<i>Isoptericola</i>	1	2	0	0	0	0	3
<i>Jonesia</i>	1	0	0	0	0	0	1
<i>Leucobacter</i>	9	0	0	0	0	0	9
<i>Microbacterium</i>	0	8	0	40	0	19	67
<i>Micrococcus</i>	2	0	0	4	3	0	9
<i>Nesterenkonia</i>	1	0	0	0	0	0	1
<i>Rhodococcus</i>	4	0	0	0	0	0	4
<i>Serinicoccus</i>	0	1	0	0	0	0	1
Phylum Bacteroidetes	7 (1.06%)	2 (0.30%)	1 (0.15%)	4 (0.61%)	0	3 (0.46%)	17 (2.58%)
Class Cytophagia	0	1 (0.15%)	0	1 (0.15%)	0	0	2 (0.31%)
<i>Algoriphagus</i>	0	0	0	1	0	0	1
<i>Marivirga</i>	0	1	0	0	0	0	1
Class Flavobacteria	7 (1.06%)	1 (0.15%)	1 (0.15%)	3 (0.46%)	0	3 (0.46%)	15 (2.29%)
<i>Flavobacterium</i>	0	0	0	0	0	1	1
<i>Joostella</i>	0	1	0	0	0	0	1
<i>Leeuwenhoekiella</i>	1	0	0	0	0	0	1
<i>Mesonina</i>	2	0	0	0	0	0	2
<i>Muricauda</i>	0	0	0	3	0	1	4
<i>Zunongwangia</i>	4	0	1	0	0	1	6

Phylum Firmicutes							
Class Bacilli	25 (3.80%)	7 (1.06%)	2 (0.31%)	10 (1.52%)	14 (2.13%)	10 (1.52%)	68 (10.33%)
<i>Aerococcus</i>	0	0	0	0	0	2	2
<i>Bacillus</i>	1	3	0	4	4	8	20
<i>Carnobacterium</i>	0	0	2	0	0	0	2
<i>Marinilactibacillus</i>	24	4	0	5	9	0	42
<i>Staphylococcus</i>	0	0	0	1	1	0	2
Phylum Proteobacteria	154 (23.40%)	61 (9.27%)	19 (2.89%)	64 (9.73%)	20 (3.04%)	105 (15.96%)	423 (64.29%)
Class Alphaproteobacteria	40 (6.12%)	35 (5.35%)	10 (1.52%)	55 (8.36%)	9 (1.37%)	79 (12.01%)	228 (34.86%)
<i>Altererythrobacter</i>	10	3	1	12	2	21	49
<i>Aquamicrobium</i>	0	0	0	0	0	2	2
<i>Aurantimonas</i>	0	0	1	1	2	3	7
<i>Breoghania</i>	0	1	0	0	0	5	6
<i>Citricella</i>	0	0	0	1	0	0	1
<i>Citreimonas</i>	1	0	0	1	0	0	2
<i>Citromicrobium</i>	2	3	0	2	1	1	9
<i>Cohaesibacter</i>	0	0	0	0	0	3	3
<i>Ensifer</i>	0	4	0	0	0	1	5
<i>Erythrobacter</i>	0	10	7	2	0	0	19
<i>Erythromicrobium</i>	0	0	0	1	0	0	1
<i>Hoeflea</i>	0	0	0	5	0	12	17
<i>Labrenzia</i>	0	6	0	4	0	12	22
<i>Loktanella</i>	0	0	0	1	0	0	1
<i>Maribius</i>	0	0	1	0	0	0	1
<i>Martelella</i>	0	1	0	2	1	3	7
<i>Novosphingobium</i>	1	0	0	0	0	0	1
<i>Oceanibulbus</i>	2	0	0	0	1	0	3
<i>Oceanicola</i>	5	2	0	0	0	1	8
<i>Oceaniovalibus</i>	1	0	0	0	0	0	1
<i>Paracoccus</i>	0	0	0	0	1	0	1
<i>Pararhodobacter</i>	0	1	0	1	0	0	2
<i>Parasphingopyxis</i>	0	0	0	1	0	0	1

<i>Rahnella</i>	0	0	0	7	0	1	8
<i>Rhizobium</i>	0	0	0	0	1	6	7
<i>Rhodovulum</i>	0	0	0	3	0	1	4
<i>Roseisalinus</i>	2	0	0	0	0	0	2
<i>Roseivivax</i>	4	1	0	0	0	0	5
<i>Roseovarius</i>	1	0	0	2	0	0	3
<i>Sphingorabdus</i>	2	0	0	1	0	1	4
<i>Stakelama</i>	4	1	0	7	0	1	13
<i>Sulfitobacter</i>	0	0	0	1	0	2	3
<i>Thalassospira</i>	2	1	0	0	0	3	6
<i>Thioclava</i>	3	0	0	0	0	0	3
<i>Yangia</i>	0	1	0	0	0	0	1
Class Betaproteobacteria	1 (0.15%)	0	0	0	0	0	1 (0.15%)
<i>Achromobacter</i>	1	0	0	0	0	0	1
Class Gammaproteobacteria	112 (17.13%)	23 (3.52%)	9 (1.37%)	9 (1.37%)	11 (1.67%)	26 (3.95%)	190 (29.05%)
<i>Chromohalobacter</i>	2	0	0	0	1	0	3
<i>Gilvimarinus</i>	3	0	0	0	0	0	3
<i>Halomonas</i>	13	5	0	0	4	2	24
<i>Idiomarina</i>	5	0	0	0	0	0	5
<i>Kushneria</i>	6	0	0	0	0	0	6
<i>Marinobacter</i>	0	2	0	3	0	0	5
<i>Marinomonas</i>	0	0	2	0	0	1	3
<i>Marinospirillum</i>	7	1	1	0	0	0	9
<i>Moraxella</i>	0	0	0	0	1	0	1
<i>Pseudoalteromonas</i>	1	1	0	0	0	0	2
<i>Pseudomonas</i>	1	0	0	5	4	5	15
<i>Pseudoxanthomonas</i>	1	0	0	0	0	0	1
<i>Psychrobacter</i>	0	1	0	0	0	0	1
<i>Saccharospirillum</i>	1	0	0	0	0	0	1
<i>Salinicola</i>	42	16	0	0	0	0	58
<i>Shewanella</i>	0	0	0	0	0	6	6
<i>Vibrio</i>	31	0	6	1	1	12	51

AG, aboveground; BG, belowground.

**Supplementary Table S3.3** Taxonomic attribution, accession number and results for enzymatic activity and plant growth promotion assays of 467 representative isolates.

ID	Class	Genus	Number of isolates in group	Accession number	Enzymatic activity					Plant growth promotion traits						
					Amylolytic	Proteolytic	Lipolytic	Cellulolytic	Xylanolytic	Pectinolytic (pH 5.0)	Pectinolytic (pH 7.0)	Phosphate solubilization	IAA production	Siderophore production	Presence of <i>nifH</i> gene	ACC deaminase
CAG-1	g-Proteobacteria	<i>Saccharospirillum</i>	1	KT324749	-	+	nd	-	+	nd	nd	-	48.9	nd	-	-
CAG-2	a-Proteobacteria	<i>Oceanicola</i>	1	KT324750	-	-	-	-	+	nd	nd	-	17.8	nd	-	-
CAG-3	a-Proteobacteria	<i>Oceanicola</i>	3	KT324751	-	-	+	+	+	nd	nd	-	92.6	-	-	-
CAG-4	a-Proteobacteria	<i>Thalassospira</i>	1	KT324752	-	-	-	-	+	nd	nd	-	37.9	nd	-	-
CAG-5	a-Proteobacteria	<i>Novosphingobium</i>	1	KT324753	-	-	+	-	+	nd	nd	-	45.5	-	-	-
CAG-6	a-Proteobacteria	<i>Thalassospira</i>	1	KT324754	-	-	-	-	+	nd	nd	-	86.5	-	-	-
CAG-7	g-Proteobacteria	<i>Idiomarina</i>	2	KT324755	-	+	+	-	+	nd	nd	-	24.7	nd	-	-
CAG-8	Flavobacteria	<i>Zunongwangia</i>	1	KT324758	-	-	-	+	-	nd	nd	+	17.8	nd	-	+
CAG-9	g-Proteobacteria	<i>Vibrio</i>	1	KT324759	+	+	+	+	+	nd	nd	-	33.1	nd	-	-
CAG-10	a-Proteobacteria	<i>Stakelama</i>	2	KT324760	+	-	+	+	+	nd	-	-	11.3	-	-	+
CAG-11	a-Proteobacteria	<i>Roseivivax</i>	1	KT324761	-	-	-	-	+	nd	nd	-	49.7	+	+	-
CAG-12	g-Proteobacteria	<i>Vibrio</i>	1	KT324762	-	+	+	+	-	nd	nd	-	29.2	-	-	-
CAG-13	g-Proteobacteria	<i>Marinospirillum</i>	3	KT324763	-	+	-	nd	+	nd	nd	-	16.6	-	-	-
CAG-14	Flavobacteria	<i>Mesonina</i>	1	KT324764	-	nd	-	nd	-	nd	nd	-	48.3	nd	-	-
CAG-15	g-Proteobacteria	<i>Marinospirillum</i>	1	KT324765	-	+	+	-	+	nd	nd	-	48.6	+	-	-
CAG-16	g-Proteobacteria	<i>Marinospirillum</i>	1	KT324766	+	+	-	-	-	nd	nd	-	32.2	-	-	-
CAG-17	g-Proteobacteria	<i>Vibrio</i>	3	KT324767	+	+	+	+	+	nd	-	-	2.0	-	+	-
CAG-18	a-Proteobacteria	<i>Oceanibulbus</i>	1	KT324768	+	-	+	+	+	-	-	+	41.4	+	+	+
CAG-19	g-Proteobacteria	<i>Salinicola</i>	4	KT324769	+	-	+	+	+	-	-	+	129.5	-	-	+
CAG-20	a-Proteobacteria	<i>Stakelama</i>	2	KT324770	-	-	-	-	-	nd	-	-	44.5	-	-	-
CAG-21	g-Proteobacteria	<i>Salinicola</i>	4	KT324771	+	-	+	-	+	-	-	+	34.8	-	-	+
CAG-22	Bacilli	<i>Marinilactibacillus</i>	1	KT324772	-	+	-	-	+	nd	-	-	37.3	+	-	-
CAG-23	Bacilli	<i>Marinilactibacillus</i>	1	KT324773	-	+	-	-	+	nd	-	-	19.3	+	-	-
CAG-24	Bacilli	<i>Marinilactibacillus</i>	1	KT324774	-	+	-	-	+	nd	-	-	4.2	+	-	-

CAG-25	Bacilli	<i>Marinilactibacillus</i>	1	KT324775	-	+	-	-	+	nd	-	-	9.6	+	-	-
CAG-26	Bacilli	<i>Marinilactibacillus</i>	2	KT324776	-	+	-	-	+	nd	-	-	-0.3	+	-	-
CAG-27	Bacilli	<i>Marinilactibacillus</i>	1	KT324777	-	+	-	-	+	nd	-	-	5.1	+	-	-
CAG-28	Bacilli	<i>Marinilactibacillus</i>	1	KT324778	-	+	-	+	+	nd	-	-	12.4	+	-	-
CAG-29	Bacilli	<i>Marinilactibacillus</i>	1	KT324779	+	-	+	+	+	nd	+	-	2.6	-	-	+
CAG-30	g-Proteobacteria	<i>Kushneria</i>	2	KT324780	+	-	+	+	+	nd	+	-	5.0	-	-	+
CAG-31	g-Proteobacteria	<i>Chromohalobacter</i>	2	KT324781	+	+	+	+	+	nd	+	-	3.6	+	-	+
CAG-32	g-Proteobacteria	<i>Halomonas</i>	1	KT324782	nd	nd	nd	nd	nd	nd	nd	-	4.9	+	-	+
CAG-33	Bacilli	<i>Marinilactibacillus</i>	1	KT324783	+	+	-	+	+	nd	-	-	50.2	+	-	-
CAG-34	Bacilli	<i>Marinilactibacillus</i>	1	KT324784	-	+	-	+	+	nd	-	-	41.5	+	-	-
CAG-35	Bacilli	<i>Marinilactibacillus</i>	1	KT324785	-	+	-	+	+	nd	-	-	30.3	+	-	-
CAG-36	a-Proteobacteria	<i>Oceanibulbus</i>	1	KT324786	+	-	-	-	-	nd	nd	+	55.7	nd	+	+
CAG-37	g-Proteobacteria	<i>Kushneria</i>	1	KT324787	-	-	-	-	-	nd	nd	-	58.4	nd	-	-
CAG-38	Bacilli	<i>Marinilactibacillus</i>	1	KT324788	+	+	-	+	+	nd	-	-	72.9	-	-	-
CAG-39	Bacilli	<i>Marinilactibacillus</i>	1	KT324789	+	+	-	+	+	nd	-	-	43.5	-	-	-
CAG-40	g-Proteobacteria	<i>Halomonas</i>	3	KT324790	+	+	+	-	+	nd	nd	-	94.3	-	-	-
CAG-41	a-Proteobacteria	<i>Roseivivax</i>	1	KT324791	+	-	-	+	+	nd	nd	-	108.3	-	+	-
CAG-42	g-Proteobacteria	<i>Halomonas</i>	1	KT324792	+	-	+	-	+	nd	nd	-	84.0	-	-	-
CAG-43	a-Proteobacteria	<i>Roseivivax</i>	1	KT324793	-	-	+	-	-	nd	nd	-	33.3	nd	+	-
CAG-44	a-Proteobacteria	<i>Citromicrobium</i>	1	KT324794	-	-	+	-	-	nd	nd	-	39.1	-	-	-
CAG-45	a-Proteobacteria	<i>Roseivivax</i>	1	KT324795	-	-	-	-	+	nd	nd	-	27.8	nd	+	-
CAG-46	a-Proteobacteria	<i>Altererythrobacter</i>	2	KT324796	-	-	+	+	+	nd	nd	-	13.0	-	-	-
CAG-47	a-Proteobacteria	<i>Roseisalinus</i>	1	KT324797	-	-	+	-	-	nd	nd	-	35.0	-	-	-
CAG-48	Bacilli	<i>Marinilactibacillus</i>	1	KT324798	-	+	-	+	+	nd	-	-	21.0	+	-	-
CAG-49	a-Proteobacteria	<i>Roseisalinus</i>	1	KT324799	-	-	+	-	-	nd	nd	-	30.3	-	-	-
CAG-50	g-Proteobacteria	<i>Idiomarina</i>	1	KT324800	-	+	-	-	+	nd	nd	-	35.9	+	-	-
CAG-51	g-Proteobacteria	<i>Halomonas</i>	1	KT324801	-	-	+	-	+	nd	nd	-	41.9	-	-	-
CAG-52	Bacilli	<i>Marinilactibacillus</i>	1	KT324802	-	+	-	+	+	nd	-	-	9.2	-	-	-
CAG-53	a-Proteobacteria	<i>Citromicrobium</i>	1	KT324803	-	-	+	-	-	nd	nd	-	44.7	-	+	-
CAG-54	a-Proteobacteria	<i>Thioclava</i>	1	KT324815	-	-	-	-	+	nd	nd	-	62.6	-	-	+
CAG-55	a-Proteobacteria	<i>Oceanicola</i>	1	KT324816	-	-	nd	-	+	nd	nd	-	83.4	nd	-	+
CAG-56	g-Proteobacteria	<i>Vibrio</i>	1	KT324817	+	+	-	-	+	-	+	-	6.2	+	-	-
CAG-57	g-Proteobacteria	<i>Salinicola</i>	1	KT324818	-	-	+	-	-	nd	nd	+	79.7	-	-	+
CAG-58	Flavobacteria	<i>Zunongwangia</i>	1	KT324819	-	-	+	-	+	nd	nd	+	76.1	-	-	-
CAG-59	Flavobacteria	<i>Zunongwangia</i>	1	KT324820	nd	nd	nd	nd	nd	nd	nd	-	31.4	nd	-	-
CAG-60	g-Proteobacteria	<i>Salinicola</i>	1	KT324821	+	-	+	+	-	+	+	+	4.2	+	-	+
CAG-61	g-Proteobacteria	<i>Halomonas</i>	1	KT324822	-	-	+	-	+	nd	nd	-	34.2	+	-	-
CAG-62	g-Proteobacteria	<i>Salinicola</i>	1	KT324823	-	-	+	-	+	nd	nd	+	79.2	nd	-	-
CAG-63	g-Proteobacteria	<i>Salinicola</i>	1	KT324824	+	+	-	-	+	nd	+	-	2.0	+	-	+

CAG-64	g-Proteobacteria	<i>Salinicola</i>	1	KT324825	+	-	+	-	+	-	+	+	33.8	-	-	+
CAG-65	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324826	+	-	+	-	-	nd	-	-	44.7	-	-	+
CAG-66	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324827	+	-	+	-	+	-	+	+	5.5	-	-	+
CAG-67	Actinobacteria_c	<i>Curtobacterium</i>	1	KT324828	+	+	-	+	-	-	-	-	4.3	-	-	+
CAG-68	Actinobacteria_c	<i>Demequina</i>	1	KT324829	-	+	+	+	+	nd	nd	-	51.4	nd	-	-
CAG-69	Bacilli	<i>Marinilactibacillus</i>	1	KT324830	-	+	+	-	+	nd	nd	-	36.7	-	-	-
CAG-70	g-Proteobacteria	<i>Pseudoxanthomonas</i>	1	KT324831	+	-	+	+	+	nd	+	-	19.0	-	-	+
CAG-71	g-Proteobacteria	<i>Salinicola</i>	1	KT324832	+	-	+	-	+	-	+	+	36.9	+	-	+
CAG-72	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324833	+	-	+	+	+	-	+	-	5.5	-	-	+
CAG-73	Actinobacteria_c	<i>Curtobacterium</i>	2	KT324834	-	+	-	+	+	-	-	-	7.3	+	-	+
CAG-74	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324835	-	-	+	-	-	nd	-	-	67.5	-	-	+
CAG-75	g-Proteobacteria	<i>Salinicola</i>	16	KT324836	+	-	+	-	+	-	+	+	42.7	-	-	+
CAG-76	Bacilli	<i>Marinilactibacillus</i>	1	KT324837	-	+	+	-	+	nd	nd	-	46.1	-	-	+
CAG-77	Bacilli	<i>Marinilactibacillus</i>	1	KT324838	-	+	-	-	+	nd	nd	+	101.7	-	-	+
CAG-78	a-Proteobacteria	<i>Altererythrobacter</i>	2	KT324839	+	-	+	-	-	nd	-	-	33.1	-	-	+
CAG-79	g-Proteobacteria	<i>Salinicola</i>	1	KT324840	+	-	+	-	+	-	+	+	27.8	+	-	+
CAG-80	g-Proteobacteria	<i>Vibrio</i>	2	KT324841	+	+	-	+	+	nd	+	-	43.8	+	-	-
CAG-81	g-Proteobacteria	<i>Vibrio</i>	1	KT324842	-	-	+	+	+	nd	nd	+	75.6	+	-	+
CAG-82	Actinobacteria_c	<i>Curtobacterium</i>	2	KT324843	-	+	-	+	+	-	-	-	8.9	+	-	+
CAG-83	g-Proteobacteria	<i>Kushneria</i>	1	KT324844	+	-	-	-	+	-	+	+	5.1	+	-	+
CAG-84	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324845	-	-	+	-	-	nd	-	-	0.7	-	-	+
CAG-85	g-Proteobacteria	<i>Vibrio</i>	1	KT324846	+	+	-	+	-	nd	-	-	4.5	+	+	-
CAG-86	g-Proteobacteria	<i>Vibrio</i>	1	KT324847	+	+	-	+	+	nd	+	+	24.5	+	+	+
CAG-87	g-Proteobacteria	<i>Vibrio</i>	1	KT324848	+	+	-	+	+	nd	+	-	22.8	+	-	-
CAG-88	g-Proteobacteria	<i>Vibrio</i>	1	KT324849	+	+	-	+	+	nd	-	-	37.5	+	-	-
CAG-89	Actinobacteria_c	<i>Demequina</i>	1	KT324850	-	-	+	+	+	+	+	-	53.7	-	-	-
CAG-90	g-Proteobacteria	<i>Salinicola</i>	7	KT324851	+	-	+	+	+	-	+	+	24.8	+	-	+
CAG-91	g-Proteobacteria	<i>Vibrio</i>	3	KT324852	+	+	-	+	+	+	+	-	33.5	+	-	+
CAG-92	g-Proteobacteria	<i>Salinicola</i>	12	KT324853	+	-	+	+	+	-	-	+	29.7	+	-	+
CAG-93	g-Proteobacteria	<i>Halomonas</i>	1	KT324899	+	-	+	+	+	nd	+	-	3.6	+	-	+
CAG-94	Actinobacteria_c	<i>Leucobacter</i>	1	KT324900	-	-	-	+	+	nd	-	-	9.1	+	-	+
CAG-95	Actinobacteria_c	<i>Nesterenkonia</i>	1	KT324901	+	-	nd	+	+	nd	+	-	2.5	-	-	+
CAG-96	Actinobacteria_c	<i>Dietzia</i>	1	KT324902	+	-	nd	+	+	nd	nd	-	2.6	-	-	-
CAG-97	Actinobacteria_c	<i>Dietzia</i>	1	KT324903	+	-	nd	+	+	nd	+	-	8.2	-	-	+
CAG-98	nd	nd	1	nd	+	-	nd	-	+	nd	+	-	5.4	-	-	+
CAG-99	Actinobacteria_c	<i>Dietzia</i>	6	KT324904	+	-	nd	+	+	nd	+	-	7.8	-	-	+
CAG-100	g-Proteobacteria	<i>Halomonas</i>	1	KT324905	-	-	nd	-	-	nd	nd	-	72.1	-	-	+
CAG-101	g-Proteobacteria	<i>Kushneria</i>	2	KT324906	+	-	-	-	-	nd	+	+	3.0	-	-	+
CAG-102	b-Proteobacteria	<i>Achromobacter</i>	1	KT324907	-	+	nd	+	+	nd	-	-	9.2	-	-	+

CAG-103	Bacilli	<i>Marinilactibacillus</i>	1	KT324908	-	-	nd	+	+	nd	+	-	37.9	-	-	+
CAG-104	nd	nd	1	nd	+	+	-	-	+	+	+	-	3.0	+	-	+
CAG-105	Actinobacteria_c	<i>Leucobacter</i>	1	KT324909	-	-	-	-	+	nd	-	-	-0.3	+	-	+
CAG-106	Actinobacteria_c	<i>Rhodococcus</i>	4	KT324910	+	-	nd	-	-	nd	+	-	2.6	+	-	+
CAG-107	Bacilli	<i>Marinilactibacillus</i>	1	KT324911	-	-	nd	+	+	nd	+	-	8.6	-	-	+
CAG-108	Actinobacteria_c	<i>Leucobacter</i>	7	KT324912	+	-	-	+	+	nd	-	-	7.3	+	-	-
CAG-109	g-Proteobacteria	<i>Halomonas</i>	1	KT324913	+	-	+	+	+	nd	+	-	8.9	+	-	+
CAG-110	Bacilli	<i>Marinilactibacillus</i>	1	KT324914	-	-	nd	+	+	nd	-	-	0.9	-	-	-
CAG-111	Bacilli	<i>Marinilactibacillus</i>	1	KT324915	-	-	nd	+	+	nd	-	-	13.1	-	-	-
CAG-112	g-Proteobacteria	<i>Idiomarina</i>	1	KT324916	-	+	+	+	+	nd	nd	-	4.6	-	-	+
CAG-113	g-Proteobacteria	<i>Halomonas</i>	1	KT324917	-	-	+	+	+	nd	+	-	13.0	-	-	-
CAG-114	Actinobacteria_c	<i>Dietzia</i>	2	KT324918	+	+	-	+	+	+	+	-	4.9	-	-	+
CAG-115	Actinobacteria_c	<i>Frigoribacterium</i>	1	KT324919	-	-	nd	+	-	nd	+	-	11.3	-	-	-
CAG-116	Actinobacteria_c	<i>Dietzia</i>	1	KT324920	+	-	nd	+	+	nd	+	-	6.4	-	-	+
CAG-117	Bacilli	<i>Marinilactibacillus</i>	1	KT324921	-	-	nd	+	+	nd	+	-	7.3	+	-	-
CAG-118	a-Proteobacteria	<i>Thioclava</i>	1	KT324922	-	-	+	-	-	nd	nd	-	72.7	-	-	-
CAG-119	a-Proteobacteria	<i>Sphingorabdus</i>	2	KT324923	-	-	+	-	-	nd	nd	-	46.1	nd	-	-
CAG-120	a-Proteobacteria	<i>Citireimonas</i>	1	KT324924	-	+	+	+	+	nd	nd	-	50.4	-	-	-
CAG-121	g-Proteobacteria	<i>Vibrio</i>	4	KT324925	+	+	+	+	+	nd	-	-	44.6	+	-	+
CAG-122	g-Proteobacteria	<i>Vibrio</i>	1	KT324926	+	+	+	+	+	nd	+	-	41.2	-	-	+
CAG-123	g-Proteobacteria	<i>Vibrio</i>	1	KT324927	+	+	+	+	+	nd	-	-	40.0	+	+	+
CAG-124	g-Proteobacteria	<i>Vibrio</i>	2	KT324928	+	+	+	+	+	nd	+	-	35.0	+	+	+
CAG-125	Actinobacteria_c	<i>Micrococcus</i>	1	KT324929	+	+	+	+	+	nd	+	-	42.0	-	+	+
CAG-126	g-Proteobacteria	<i>Vibrio</i>	1	KT324930	+	+	+	+	+	nd	+	-	41.7	+	-	+
CAG-127	g-Proteobacteria	<i>Halomonas</i>	1	KT324931	+	+	+	+	+	nd	+	-	57.9	-	+	-
CAG-128	g-Proteobacteria	<i>Marinospirillum</i>	1	KT324932	+	+	+	+	+	nd	+	-	36.0	+	+	-
CAG-129	g-Proteobacteria	<i>Marinospirillum</i>	1	KT324933	-	+	+	-	-	nd	nd	-	46.3	-	-	-
CAG-130	g-Proteobacteria	<i>Vibrio</i>	1	KT324934	+	+	+	+	+	nd	+	-	58.0	+	+	-
CAG-131	g-Proteobacteria	<i>Vibrio</i>	1	KT324935	+	+	+	+	+	nd	+	+	57.8	+	+	+
CAG-132	g-Proteobacteria	<i>Vibrio</i>	4	KT324936	+	+	+	+	+	nd	+	-	51.6	+	+	-
CAG-133	Flavobacteria	<i>Mesonina</i>	1	KT324937	-	+	+	-	-	nd	nd	-	19.0	-	-	-
CAG-134	g-Proteobacteria	<i>Halomonas</i>	1	KT324938	-	+	-	+	+	nd	nd	-	26.9	-	-	-
CAG-135	Actinobacteria_c	<i>Demequina</i>	1	KT324939	-	+	-	+	+	nd	nd	-	15.8	nd	-	-
CAG-136	g-Proteobacteria	<i>Gilvimarinus</i>	1	KT324940	+	+	+	+	+	nd	nd	-	22.3	-	-	-
CAG-137	a-Proteobacteria	<i>Roseovarius</i>	1	KT324941	-	+	-	+	+	nd	nd	-	57.0	-	+	-
CAG-138	a-Proteobacteria	<i>Oceaniovalibus</i>	1	KT324942	-	-	-	+	+	nd	nd	-	54.1	-	+	-
CAG-139	g-Proteobacteria	<i>Pseudoalteromonas</i>	1	KT324943	-	+	+	+	+	nd	nd	-	23.0	-	-	-
CAG-140	g-Proteobacteria	<i>Idiomarina</i>	1	KT324944	-	+	+	+	+	nd	nd	-	34.5	-	-	-
CAG-141	Flavobacteria	<i>Zunongwangia</i>	1	KT324945	nd	nd	-	nd	nd	nd	nd	-	23.8	-	-	-

CAG-142	a-Proteobacteria	<i>Thioclava</i>	1	KT324946	-	-	-	+	+	nd	nd	-	55.1	nd	-	-
CAG-143	g-Proteobacteria	<i>Gilvimarinus</i>	2	KT324947	+	+	+	+	+	nd	nd	-	46.1	nd	-	-
CAG-144	Flavobacteria	<i>Leeuwenhoekella</i>	1	KT324948	+	-	+	+	+	nd	nd	-	48.2	nd	-	-
CAG-145	g-Proteobacteria	<i>Pseudomonas</i>	1	KT324949	+	-	+	+	-	nd	nd	-	34.1	-	-	-
CAG-146	Actinobacteria_c	<i>Isoptericola</i>	1	KT324950	+	-	-	+	+	nd	nd	-	40.7	nd	-	-
CAG-147	g-Proteobacteria	<i>Salinicola</i>	1	KT324951	-	-	+	-	-	-	+	-	4.6	-	-	-
CAG-148	Bacilli	<i>Bacillus</i>	1	KT324952	-	-	-	+	+	nd	-	-	14.4	-	-	+
CAG-149	Actinobacteria_c	<i>Micrococcus</i>	1	KT324953	+	+	-	+	+	nd	+	-	29.3	+	-	+
CAG-150	Actinobacteria_c	<i>Jonesia</i>	1	KT324954	-	-	+	+	+	nd	nd	-	62.1	+	-	+
CBG-1	Actinobacteria_c	<i>Isoptericola</i>	1	KT324756	+	+	+	+	+	nd	nd	-	58.2	-	-	+
CBG-2	nd	nd	1	nd	-	+	-	+	+	nd	nd	-	36.5	nd	-	+
CBG-3	Actinobacteria_c	<i>Isoptericola</i>	1	KT324757	+	+	+	+	+	nd	nd	-	96.5	-	-	+
CBG-4	Bacilli	<i>Marinilactibacillus</i>	1	KT324804	-	-	+	nd	+	nd	nd	-	52.6	-	-	-
CBG-5	a-Proteobacteria	<i>Labrenzia</i>	1	KT324805	-	-	-	-	+	nd	nd	-	34.0	-	-	+
CBG-6	Bacilli	<i>Marinilactibacillus</i>	1	KT324806	-	-	-	nd	+	nd	nd	-	37.3	-	-	+
CBG-7	Bacilli	<i>Marinilactibacillus</i>	1	KT324807	-	-	-	+	+	nd	nd	-	41.0	-	-	+
CBG-8	g-Proteobacteria	<i>Marinobacter</i>	1	KT324808	+	-	+	+	+	nd	nd	-	43.5	-	-	+
CBG-9	Bacilli	<i>Bacillus</i>	1	KT324809	+	+	+	-	-	nd	nd	-	12.1	-	-	+
CBG-10	a-Proteobacteria	<i>Roseivivax</i>	1	KT324810	-	-	-	-	-	nd	nd	-	25.4	nd	-	-
CBG-11	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324811	-	-	+	-	+	nd	nd	-	37.8	nd	-	-
CBG-12	a-Proteobacteria	<i>Pararhodobacter</i>	1	KT324812	+	-	nd	+	nd	+	+	-	45.9	-	-	+
CBG-13	Bacilli	<i>Bacillus</i>	1	KT324813	+	+	+	-	+	nd	nd	-	10.2	-	-	-
CBG-14	Bacilli	<i>Marinilactibacillus</i>	1	KT324814	-	+	-	+	+	nd	-	-	31.7	+	-	+
CBG-15	Bacilli	<i>Bacillus</i>	1	KT324854	+	+	+	-	+	nd	nd	-	15.0	+	-	-
CBG-16	a-Proteobacteria	<i>Yangia</i>	1	KT324855	-	-	-	-	-	nd	nd	-	23.1	+	-	-
CBG-17	Actinobacteria_c	<i>Demequina</i>	1	KT324856	-	+	+	+	+	nd	nd	-	50.6	nd	-	-
CBG-18	a-Proteobacteria	<i>Breoghania</i>	2	KT324857	-	-	+	+	-	nd	-	-	32.2	-	+	+
CBG-19	Actinobacteria_c	<i>Serinicoccus</i>	1	KT324858	+	+	+	+	-	nd	nd	-	42.2	-	-	+
CBG-20	a-Proteobacteria	<i>Martellella</i>	1	KT324859	-	+	-	-	-	nd	nd	-	56.2	-	-	-
CBG-21	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324860	+	-	+	+	+	nd	nd	-	54.5	nd	-	-
CBG-22	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324861	-	-	+	-	+	nd	nd	-	32.7	-	-	-
CBG-23	Cytophagia	<i>Marivirga</i>	1	KT324862	+	+	+	+	+	nd	nd	-	50.3	nd	-	-
CBG-24	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324863	-	-	+	-	+	nd	nd	-	33.3	nd	-	-
CBG-25	a-Proteobacteria	<i>Citromicrobium</i>	1	KT324864	-	+	+	-	+	nd	nd	-	64.2	-	-	-
CBG-26	a-Proteobacteria	<i>Ensifer</i>	1	KT324865	-	-	+	-	-	+	nd	-	71.8	-	-	+
CBG-27	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324866	+	-	-	-	+	nd	+	-	47.1	-	-	+
CBG-28	a-Proteobacteria	<i>Ensifer</i>	1	KT324867	-	+	-	-	-	-	-	-	70.6	+	-	+
CBG-29	Actinobacteria_c	<i>Demequina</i>	1	KT324868	-	+	-	+	+	nd	nd	-	45.0	-	-	-
CBG-30	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324869	+	-	+	-	-	-	-	+	22.2	-	-	+



CBG-31	a-Proteobacteria	<i>Ensifer</i>	2	KT324870	-	+	-	-	-	+	+	+	67.8	+	-	+
CBG-32	Actinobacteria_c	<i>Demequina</i>	1	KT324871	-	+	-	+	+	nd	nd	-	43.8	-	-	-
CBG-33	Actinobacteria_c	<i>Microbacterium</i>	1	KT324872	-	+	-	-	+	nd	nd	-	88.2	-	-	+
CBG-34	Actinobacteria_c	<i>Demequina</i>	1	KT324873	-	+	-	+	+	nd	nd	-	6.4	nd	-	+
CBG-35	Actinobacteria_c	<i>Microbacterium</i>	1	KT324874	+	-	-	+	-	nd	+	-	3.9	+	-	+
CBG-36	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324875	-	-	-	-	-	nd	-	-	27.3	-	-	+
CBG-37	g-Proteobacteria	<i>Halomonas</i>	1	KT324876	+	-	+	+	-	nd	-	-	57.3	+	-	+
CBG-38	g-Proteobacteria	<i>Halomonas</i>	1	KT324877	-	-	-	-	-	nd	nd	-	80.6	+	-	-
CBG-39	g-Proteobacteria	<i>Halomonas</i>	1	KT324878	+	-	+	+	-	-	+	+	2.6	+	-	+
CBG-40	a-Proteobacteria	<i>Citromicrobium</i>	1	KT324879	+	-	-	+	+	nd	-	-	2.4	+	-	+
CBG-41	a-Proteobacteria	<i>Thalassospira</i>	1	KT324880	-	-	-	-	-	nd	+	-	93.8	-	-	-
CBG-42	a-Proteobacteria	<i>Citromicrobium</i>	1	KT324881	-	-	+	-	+	nd	nd	-	48.2	-	-	-
CBG-43	g-Proteobacteria	<i>Marinobacter</i>	1	KT324882	+	+	+	-	-	nd	nd	-	21.8	-	-	-
CBG-44	a-Proteobacteria	<i>Oceanicola</i>	1	KT324883	+	+	+	-	-	nd	nd	-	22.3	-	-	-
CBG-45	g-Proteobacteria	<i>Salinicola</i>	1	KT324884	-	-	+	-	-	nd	+	+	27.3	+	-	+
CBG-46	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324885	+	-	-	+	+	nd	-	-	7.2	+	-	+
CBG-47	Actinobacteria_c	<i>Microbacterium</i>	1	KT324886	+	-	-	+	+	nd	-	-	11.9	+	-	+
CBG-48	a-Proteobacteria	<i>Stakelama</i>	1	KT324887	+	-	-	+	+	-	+	-	8.3	-	-	+
CBG-49	Actinobacteria_c	<i>Dietzia</i>	2	KT324888	+	-	-	+	+	nd	-	-	3.6	-	-	+
CBG-50	g-Proteobacteria	<i>Salinicola</i>	2	KT324889	+	-	+	+	-	-	+	+	45.7	+	-	+
CBG-51	Actinobacteria_c	<i>Dietzia</i>	1	KT324890	+	-	-	+	+	nd	-	-	4.5	+	-	+
CBG-52	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324891	+	-	+	-	+	nd	+	-	9.5	-	-	+
CBG-53	Actinobacteria_c	<i>Microbacterium</i>	2	KT324892	+	-	-	+	+	nd	+	-	15.4	+	-	+
CBG-54	g-Proteobacteria	<i>Salinicola</i>	1	KT324893	+	-	+	+	-	-	-	+	3.0	+	-	+
CBG-55	g-Proteobacteria	<i>Halomonas</i>	2	KT324894	+	-	+	-	+	nd	+	-	64.2	+	-	+
CBG-56	Actinobacteria_c	<i>Microbacterium</i>	2	KT324895	+	-	+	-	+	nd	+	-	8.1	+	-	+
CBG-57	g-Proteobacteria	<i>Salinicola</i>	3	KT324896	-	+	+	-	-	+	-	-	-0.4	+	-	+
CBG-58	Actinobacteria_c	<i>Microbacterium</i>	1	KT324897	+	-	-	-	+	nd	-	-	3.9	+	-	+
CBG-59	Actinobacteria_c	<i>Dietzia</i>	2	KT324898	-	-	+	+	-	nd	nd	-	25.1	-	-	+
CBG-60	g-Proteobacteria	<i>Salinicola</i>	1	KT324955	+	-	+	-	+	+	+	+	10.2	+	-	+
CBG-61	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324956	-	+	+	-	+	nd	nd	-	10.7	-	-	-
CBG-62	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324957	+	-	nd	+	+	nd	nd	-	41.7	-	-	-
CBG-63	Flavobacteria	<i>Joostella</i>	1	KT324958	+	+	nd	+	+	nd	+	-	14.2	-	-	-
CBG-64	a-Proteobacteria	<i>Labrenzia</i>	1	KT324959	-	+	nd	+	+	nd	-	-	59.9	-	-	+
CBG-65	a-Proteobacteria	<i>Oceanicola</i>	1	KT324960	-	nd	nd	+	+	nd	nd	-	0.5	nd	-	-
CBG-66	g-Proteobacteria	<i>Marinospirillum</i>	1	KT324961	+	+	nd	nd	-	nd	-	-	46.5	-	-	+
CBG-67	g-Proteobacteria	<i>Pseudoalteromonas</i>	1	KT324962	+	+	nd	+	+	nd	-	-	59.9	-	-	-
CBG-68	a-Proteobacteria	<i>Labrenzia</i>	1	KT324963	-	+	nd	+	+	nd	nd	-	56.7	-	-	+
CBG-69	a-Proteobacteria	<i>Labrenzia</i>	1	KT324964	-	-	nd	+	+	nd	-	-	2.2	-	+	+

CBG-70	Actinobacteria_c	<i>Demequina</i>	1	KT324965	nd	nd	nd	nd	nd	nd	nd	-	48.2	nd	-	-
CBG-71	a-Proteobacteria	<i>Erythrobacter</i>	1	KT324966	-	-	+	+	+	nd	nd	-	43.1	-	-	-
CBG-72	nd	nd	1	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	-	nd
CBG-73	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324967	-	-	nd	+	+	nd	-	-	20.1	+	-	+
CBG-74	g-Proteobacteria	<i>Psychrobacter</i>	1	KT324968	-	+	nd	-	+	nd	nd	-	44.4	+	-	+
EAG-1	g-Proteobacteria	<i>Vibrio</i>	2	KT325098	-	+	+	+	+	-	-	-	43.4	+	+	-
EAG-2	g-Proteobacteria	<i>Marinospirillum</i>	1	KT325099	+	+	+	-	-	nd	+	-	65.0	+	+	-
EAG-3	a-Proteobacteria	<i>Aurantimonas</i>	1	KT325100	-	-	-	+	-	nd	nd	-	48.3	-	+	-
EAG-4	Bacilli	<i>Carnobacterium</i>	1	KT325101	nd	nd	nd	nd	nd	nd	nd	-	9.0	-	+	-
EAG-5	g-Proteobacteria	<i>Marinomonas</i>	1	KT325102	-	-	-	+	+	nd	nd	-	53.0	nd	+	-
EAG-6	g-Proteobacteria	<i>Marinomonas</i>	1	KT325103	-	-	+	+	+	nd	nd	-	49.4	-	+	-
EAG-7	Bacilli	<i>Carnobacterium</i>	1	KT325104	nd	nd	nd	nd	nd	nd	nd	-	15.5	-	+	-
EAG-8	Flavobacteria	<i>Zunongwangia</i>	1	KT325105	+	+	+	+	-	nd	nd	-	29.7	-	+	-
EAG-9	g-Proteobacteria	<i>Vibrio</i>	3	KT325106	-	+	+	+	+	-	+	-	32.8	-	+	-
EAG-10	g-Proteobacteria	<i>Vibrio</i>	1	KT325107	-	+	+	+	+	-	-	-	37.8	-	+	-
EAG-11	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325173	-	-	+	+	-	nd	nd	-	35.4	-	-	-
EAG-12	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325174	-	-	+	+	-	nd	nd	-	54.0	nd	-	-
EAG-13	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325175	+	+	+	-	-	nd	nd	-	53.5	nd	-	-
EAG-14	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325176	+	-	+	-	+	nd	nd	-	95.0	-	-	-
EAG-15	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325177	+	-	+	-	-	nd	nd	-	68.8	-	-	-
EAG-16	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325178	-	-	+	-	-	nd	nd	-	53.0	-	-	-
EAG-17	a-Proteobacteria	<i>Maribius</i>	1	KT325179	+	nd	+	-	nd	nd	nd	-	50.2	-	+	-
EAG-18	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325180	-	-	+	-	-	nd	nd	-	27.5	nd	-	-
EBG-1	Actinobacteria_c	<i>Micrococcus</i>	1	KT325108	-	+	-	+	+	nd	-	-	12.8	-	+	+
EBG-2	Bacilli	<i>Bacillus</i>	1	KT325109	+	+	+	+	+	nd	-	-	-0.3	+	+	+
EBG-3	Bacilli	<i>Marinilactibacillus</i>	1	KT325110	-	-	-	+	+	nd	nd	-	8.8	+	+	-
EBG-4	Actinobacteria_c	<i>Demequina</i>	1	KT325111	-	-	-	+	+	nd	nd	-	47.3	-	-	-
EBG-5	Actinobacteria_c	<i>Demequina</i>	2	KT325112	-	-	-	+	+	nd	nd	-	51.3	-	+	-
EBG-6	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325113	-	-	+	+	+	nd	nd	-	8.3	nd	+	-
EBG-7	a-Proteobacteria	<i>Sphingorabdus</i>	1	KT325114	-	-	+	nd	-	nd	nd	-	7.3	nd	+	-
EBG-8	Actinobacteria_c	<i>Demequina</i>	1	KT325115	-	-	-	-	+	nd	nd	-	50.9	nd	+	-
EBG-9	Bacilli	<i>Bacillus</i>	1	KT325116	+	+	-	-	-	nd	nd	-	11.0	+	+	-
EBG-10	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325117	-	-	+	+	+	nd	nd	-	5.6	+	+	+
EBG-11	a-Proteobacteria	<i>Pararhodobacter</i>	1	KT325118	-	-	-	-	-	nd	nd	-	60.7	+	+	-
EBG-12	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325119	-	-	+	+	+	nd	nd	-	90.4	+	+	-
EBG-13	Actinobacteria_c	<i>Microbacterium</i>	2	KT325120	+	-	-	-	+	nd	nd	-	41.8	-	+	+
EBG-14	Bacilli	<i>Bacillus</i>	1	KT325121	+	+	+	+	+	nd	nd	-	63.8	-	+	+
EBG-15	a-Proteobacteria	<i>Citricella</i>	1	KT325122	-	-	-	+	-	nd	nd	-	57.9	-	+	-
EBG-16	g-Proteobacteria	<i>Marinobacter</i>	2	KT325123	+	-	+	+	+	nd	nd	-	94.4	-	+	-

EBG-17	a-Proteobacteria	<i>Roseovarius</i>	2	KT325124	-	-	-	-	-	nd	nd	-	27.8	-	+	-
EBG-18	Flavobacteria	<i>Muricauda</i>	3	KT325125	-	-	+	+	+	nd	nd	-	31.9	nd	+	-
EBG-19	Cytophagia	<i>Algoriphagus</i>	1	KT325126	-	-	-	+	+	nd	nd	-	57.7	+	+	-
EBG-20	a-Proteobacteria	<i>Rahnella</i>	8	KT325127	-	+	-	+	+	-	-	+	105.9	+	-	+
EBG-21	Actinobacteria_c	<i>Microbacterium</i>	1	KT325128	+	-	-	-	-	nd	nd	-	65.1	nd	+	+
EBG-22	Actinobacteria_c	<i>Microbacterium</i>	1	KT325129	+	-	-	+	-	nd	nd	-	41.0	nd	+	+
EBG-23	Bacilli	<i>Bacillus</i>	1	KT325130	+	+	+	-	+	nd	nd	-	43.8	+	+	+
EBG-24	Actinobacteria_c	<i>Microbacterium</i>	1	KT325131	+	-	-	+	+	nd	-	-	6.9	-	+	+
EBG-25	Bacilli	<i>Marinilactibacillus</i>	1	KT325132	-	-	-	-	+	nd	-	-	17.4	-	+	-
EBG-26	Actinobacteria_c	<i>Dietzia</i>	2	KT325133	+	+	-	+	+	nd	-	-	11.5	-	+	+
EBG-27	Bacilli	<i>Marinilactibacillus</i>	1	KT325134	-	-	-	+	+	nd	+	-	37.1	-	-	-
EBG-28	Bacilli	<i>Marinilactibacillus</i>	1	KT325135	-	-	-	+	-	nd	-	-	41.9	+	-	-
EBG-29	Actinobacteria_c	<i>Microbacterium</i>	1	KT325136	+	-	-	-	-	nd	nd	-	12.9	-	-	-
EBG-30	Actinobacteria_c	<i>Dietzia</i>	1	KT325137	+	+	-	+	+	nd	+	-	2.6	-	-	+
EBG-31	Actinobacteria_c	<i>Micrococcus</i>	1	KT325138	-	+	-	+	+	nd	+	-	14.6	+	-	+
EBG-32	g-Proteobacteria	<i>Pseudomonas</i>	2	KT325139	-	+	+	+	+	nd	nd	+	119.6	-	-	+
EBG-33	Actinobacteria_c	<i>Microbacterium</i>	1	KT325140	+	-	-	+	+	nd	-	-	7.4	-	-	+
EBG-34	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325141	+	-	+	+	+	nd	+	-	21.9	+	-	+
EBG-35	a-Proteobacteria	<i>Altererythrobacter</i>	2	KT325142	-	+	+	+	+	nd	nd	-	14.1	-	-	+
EBG-36	Actinobacteria_c	<i>Microbacterium</i>	2	KT325143	+	-	-	-	-	nd	nd	-	55.7	-	-	-
EBG-37	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325144	-	+	+	+	+	nd	nd	-	70.8	-	-	+
EBG-38	Actinobacteria_c	<i>Microbacterium</i>	1	KT325145	+	-	+	+	-	nd	nd	-	44.3	nd	-	-
EBG-39	a-Proteobacteria	<i>Erythromicrobium</i>	1	KT325146	-	+	+	-	+	nd	nd	-	26.1	-	-	-
EBG-40	Actinobacteria_c	<i>Microbacterium</i>	1	KT325147	+	-	-	-	-	nd	nd	-	51.2	nd	-	+
EBG-41	a-Proteobacteria	<i>Martellella</i>	1	KT325148	-	-	-	+	+	nd	nd	-	46.0	-	-	-
EBG-42	a-Proteobacteria	<i>Hoeflea</i>	1	KT325149	-	-	-	+	-	nd	nd	-	113.8	-	-	-
EBG-43	a-Proteobacteria	<i>Rhodovulum</i>	2	KT325150	-	-	-	-	+	nd	nd	-	49.7	-	+	+
EBG-44	g-Proteobacteria	<i>Pseudomonas</i>	2	KT325151	-	-	+	-	+	nd	nd	+	164.5	-	-	+
EBG-45	a-Proteobacteria	<i>Loktanella</i>	1	KT325152	-	-	+	-	-	nd	nd	-	28.3	nd	+	-
EBG-46	a-Proteobacteria	<i>Hoeflea</i>	1	KT325153	-	-	-	-	-	nd	nd	-	46.0	-	-	-
EBG-47	a-Proteobacteria	<i>Labrenzia</i>	3	KT325154	-	-	+	+	+	nd	nd	-	49.3	-	+	+
EBG-48	a-Proteobacteria	<i>Sulfitobacter</i>	1	KT325155	-	-	-	+	+	nd	nd	-	49.7	-	-	-
EBG-49	Actinobacteria_c	<i>Microbacterium</i>	5	KT325156	+	-	-	+	+	nd	-	+	26.2	+	-	-
EBG-50	a-Proteobacteria	<i>Hoeflea</i>	1	KT325157	-	-	-	+	-	nd	nd	-	25.7	-	+	-
EBG-51	Actinobacteria_c	<i>Micrococcus</i>	1	KT325158	-	+	-	+	+	nd	-	-	15.3	+	-	+
EBG-52	a-Proteobacteria	<i>Citireimonas</i>	1	KT325159	-	-	+	+	-	nd	nd	-	37.3	nd	-	-
EBG-53	a-Proteobacteria	<i>Martellella</i>	1	KT325160	-	-	-	+	-	nd	nd	-	13.0	-	-	-
EBG-54	g-Proteobacteria	<i>Pseudomonas</i>	1	KT325161	-	-	+	-	+	nd	nd	-	66.4	-	-	+
EBG-55	Bacilli	<i>Staphylococcus</i>	1	KT325162	-	-	-	+	+	-	+	-	5.9	+	-	-

EBG-56	Actinobacteria_c	<i>Microbacterium</i>	1	KT325163	+	-	-	+	-	nd	nd	-	44.2	-	-	+
EBG-57	Actinobacteria_c	<i>Demequina</i>	1	KT325164	+	-	-	+	+	nd	nd	-	39.8	-	-	-
EBG-58	Actinobacteria_c	<i>Microbacterium</i>	1	KT325165	+	-	+	+	-	+	+	-	12.6	-	-	-
EBG-59	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325166	+	-	+	+	+	nd	-	-	16.8	+	-	+
EBG-60	Actinobacteria_c	<i>Demequina</i>	2	KT325167	+	-	-	+	+	nd	nd	-	41.6	nd	-	-
EBG-61	Actinobacteria_c	<i>Micrococcus</i>	1	KT325168	-	-	+	+	+	nd	nd	-	59.2	+	-	+
EBG-62	Actinobacteria_c	<i>Microbacterium</i>	3	KT325169	+	-	-	+	+	nd	+	+	23.5	+	-	+
EBG-63	Bacilli	<i>Marinilactibacillus</i>	1	KT325170	+	-	-	-	+	nd	-	-	9.2	+	-	-
EBG-64	g-Proteobacteria	<i>Pseudomonas</i>	1	KT325171	-	-	+	-	-	nd	nd	+	206.4	-	-	+
EBG-65	g-Proteobacteria	<i>Vibrio</i>	1	KT325172	-	-	-	+	-	nd	nd	-	104.5	-	+	-
EBG-66	Actinobacteria_c	<i>Aeromicrobium</i>	1	KT325181	-	+	+	-	+	nd	nd	-	134.2	-	-	-
EBG-67	a-Proteobacteria	<i>Erythrobacter</i>	1	KT325182	+	-	+	+	+	nd	nd	+	23.6	+	-	+
EBG-68	a-Proteobacteria	<i>Aurantimonas</i>	1	KT325183	-	-	-	-	-	nd	nd	+	160.8	-	-	-
EBG-69	Actinobacteria_c	<i>Microbacterium</i>	1	KT325184	+	+	-	+	+	nd	-	+	19.0	+	-	+
EBG-70	a-Proteobacteria	<i>Hoeflea</i>	1	KT325185	-	-	-	-	-	nd	nd	-	134.0	+	-	-
EBG-71	Actinobacteria_c	<i>Aeromicrobium</i>	1	KT325186	+	+	+	-	-	nd	nd	-	69.7	nd	-	+
EBG-72	a-Proteobacteria	<i>Altererythrobacter</i>	2	KT325187	-	-	+	-	+	nd	nd	-	81.5	+	-	-
EBG-73	Actinobacteria_c	<i>Microbacterium</i>	1	KT325188	+	+	-	+	+	nd	-	-	15.8	-	-	+
EBG-74	a-Proteobacteria	<i>Citromicrobium</i>	2	KT325189	-	-	+	-	+	nd	nd	-	45.7	-	+	-
EBG-75	g-Proteobacteria	<i>Marinobacter</i>	1	KT325190	-	+	+	+	-	nd	nd	-	58.3	-	-	-
EBG-76	a-Proteobacteria	<i>Labrenzia</i>	2	KT325191	-	-	-	-	-	nd	nd	-	52.6	-	+	+
EBG-77	Actinobacteria_c	<i>Microbacterium</i>	2	KT325192	+	-	-	+	+	nd	+	-	5.9	-	-	+
EBG-78	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325193	-	-	+	-	-	nd	nd	-	43.6	-	-	+
EBG-79	a-Proteobacteria	<i>Stakelama</i>	2	KT325194	+	-	+	+	+	-	-	-	12.1	+	-	+
EBG-80	a-Proteobacteria	<i>Parasphingopyxis</i>	1	KT325195	-	-	+	-	+	nd	nd	-	55.9	-	-	+
EBG-81	a-Proteobacteria	<i>Hoeflea</i>	1	KT325196	-	-	-	-	-	nd	nd	-	53.6	-	+	-
EBG-82	a-Proteobacteria	<i>Stakelama</i>	5	KT325197	+	-	+	+	+	-	-	-	15.8	+	-	+
EBG-83	Actinobacteria_c	<i>Microbacterium</i>	2	KT325198	+	-	-	+	+	nd	-	-	11.0	-	-	+
EBG-84	Actinobacteria_c	<i>Microbacterium</i>	1	KT325199	+	-	+	+	+	-	-	-	6.3	-	-	+
EBG-85	Actinobacteria_c	<i>Microbacterium</i>	4	KT325200	+	-	-	-	-	nd	nd	-	49.7	-	-	-
EBG-86	Actinobacteria_c	<i>Microbacterium</i>	2	KT325201	+	+	-	+	+	nd	-	+	18.7	+	-	+
EBG-87	Actinobacteria_c	<i>Microbacterium</i>	1	KT325202	+	-	-	+	+	nd	-	-	19.0	-	-	+
EBG-88	Actinobacteria_c	<i>Microbacterium</i>	1	KT325203	+	+	-	+	+	nd	nd	+	71.4	-	-	+
EBG-89	Actinobacteria_c	<i>Microbacterium</i>	1	KT325204	+	-	-	+	+	nd	-	+	32.6	+	-	+
EBG-90	Actinobacteria_c	<i>Microbacterium</i>	1	KT325205	+	-	-	+	+	nd	-	-	6.8	+	-	+
EBG-91	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325206	+	-	+	+	+	nd	+	-	13.6	-	-	+
EBG-92	Actinobacteria_c	<i>Microbacterium</i>	1	KT325207	+	-	-	+	+	nd	-	-	-1.2	-	-	+
EBG-93	Actinobacteria_c	<i>Curtobacterium</i>	1	KT325208	+	+	+	+	+	-	-	-	6.8	+	-	+
EBG-94	Actinobacteria_c	<i>Microbacterium</i>	1	KT325209	+	-	-	+	+	nd	+	-	5.7	+	-	+

BAG-1	Actinobacteria_c	<i>Demequina</i>	2	KT324969	nd	nd	nd	nd	nd	nd	nd	-	53.2	nd	-	-
BAG-2	g-Proteobacteria	<i>Chromohalobacter</i>	1	KT324970	-	-	nd	+	-	nd	nd	-	47.3	-	-	+
BAG-3	Bacilli	<i>Bacillus</i>	1	KT324971	-	+	nd	+	+	nd	-	-	7.1	+	-	+
BAG-4	g-Proteobacteria	<i>Pseudomonas</i>	1	KT324972	-	+	+	-	-	nd	nd	+	117.3	-	-	+
BAG-5	g-Proteobacteria	<i>Halomonas</i>	5	KT324973	-	-	nd	-	+	nd	+	-	12.7	+	-	+
BAG-6	g-Proteobacteria	<i>Pseudomonas</i>	1	KT324974	+	-	nd	-	-	nd	-	+	105.0	+	-	+
BAG-7	Actinobacteria_c	<i>Micrococcus</i>	1	KT325010	-	-	-	+	+	nd	-	-	18.1	+	-	+
BAG-8	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325011	-	-	+	+	+	nd	-	-	20.7	+	-	+
BAG-9	a-Proteobacteria	<i>Oceanibulbus</i>	1	KT325012	-	-	+	-	+	nd	nd	-	59.7	nd	+	-
BAG-10	a-Proteobacteria	<i>Aurantimonas</i>	2	KT325013	-	-	-	-	+	+	nd	+	21.1	-	+	+
BAG-11	a-Proteobacteria	<i>Citromicrobium</i>	1	KT325014	-	-	+	+	+	nd	nd	-	59.7	-	-	-
BAG-12	a-Proteobacteria	<i>Martellella</i>	1	KT325015	-	-	-	-	+	nd	nd	-	59.3	-	-	+
BAG-13	Bacilli	<i>Bacillus</i>	1	KT325016	+	+	-	+	+	nd	+	-	8.8	+	-	+
BAG-14	g-Proteobacteria	<i>Moraxella</i>	1	KT325017	+	-	+	+	+	nd	-	-	10.9	+	-	-
BAG-15	Actinobacteria_c	<i>Micrococcus</i>	1	KT325018	-	+	-	+	+	nd	+	-	17.8	+	-	+
BAG-16	Actinobacteria_c	<i>Micrococcus</i>	1	KT325019	-	+	-	+	+	nd	+	-	13.7	+	-	+
BAG-17	Bacilli	<i>Staphylococcus</i>	1	KT325020	-	+	-	+	-	+	-	-	5.0	+	-	+
BAG-18	a-Proteobacteria	<i>Paracoccus</i>	1	KT325055	+	-	-	nd	nd	nd	nd	-	6.4	+	+	-
BAG-19	Bacilli	<i>Marinilactibacillus</i>	1	KT325056	-	-	-	-	+	nd	-	-	44.2	nd	-	-
BAG-20	Bacilli	<i>Marinilactibacillus</i>	1	KT325057	+	-	+	-	+	nd	nd	-	42.6	-	-	-
BAG-21	g-Proteobacteria	<i>Vibrio</i>	1	KT325058	-	+	+	+	-	+	nd	-	90.2	nd	-	-
BAG-22	Bacilli	<i>Marinilactibacillus</i>	1	KT325059	-	-	+	+	+	nd	-	-	9.2	+	-	+
BAG-23	Bacilli	<i>Marinilactibacillus</i>	1	KT325060	-	-	-	+	+	nd	-	-	22.0	-	-	-
BAG-24	Bacilli	<i>Marinilactibacillus</i>	1	KT325061	-	-	+	-	-	nd	nd	-	23.1	-	-	+
BAG-25	g-Proteobacteria	<i>Pseudomonas</i>	1	KT325062	-	-	+	-	+	nd	+	+	93.7	-	-	+
BAG-26	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325063	+	-	+	-	+	nd	+	-	8.8	-	-	+
BAG-27	Bacilli	<i>Bacillus</i>	1	KT325064	-	+	-	+	+	-	+	-	10.8	+	-	+
BAG-28	a-Proteobacteria	<i>Rhizobium</i>	2	KT325065	-	-	-	+	+	nd	nd	+	87.6	nd	-	-
BAG-29	Bacilli	<i>Bacillus</i>	1	KT325066	-	-	-	-	+	nd	-	-	10.2	-	-	-
BAG-30	Bacilli	<i>Marinilactibacillus</i>	1	KT325067	-	-	-	-	+	nd	-	-	36.4	-	-	-
BAG-31	Bacilli	<i>Marinilactibacillus</i>	1	KT325068	-	-	-	+	+	nd	-	-	8.7	+	-	+
BAG-32	Bacilli	<i>Marinilactibacillus</i>	1	KT325069	+	-	-	+	+	nd	-	-	15.8	-	-	-
BAG-33	Bacilli	<i>Marinilactibacillus</i>	1	KT325070	+	-	-	+	+	nd	-	-	10.2	+	-	-
BBG-1	a-Proteobacteria	<i>Rhizobium</i>	1	KT324975	-	-	nd	nd	-	nd	nd	-	68.8	nd	-	-
BBG-2	a-Proteobacteria	<i>Cohaesibacter</i>	1	KT324976	nd	nd	nd	nd	-	nd	nd	-	51.9	nd	-	-
BBG-3	g-Proteobacteria	<i>Shewanella</i>	3	KT324977	nd	nd	nd	nd	nd	nd	nd	-	2.6	-	-	-
BBG-4	g-Proteobacteria	<i>Shewanella</i>	1	KT324978	nd	+	nd	nd	+	nd	nd	-	14.9	-	-	-
BBG-5	a-Proteobacteria	<i>Breoghania</i>	2	KT324979	-	nd	nd	nd	+	nd	nd	-	5.8	-	-	+
BBG-6	a-Proteobacteria	<i>Thalassospira</i>	1	KT324980	-	-	nd	-	-	nd	nd	-	44.0	-	-	+

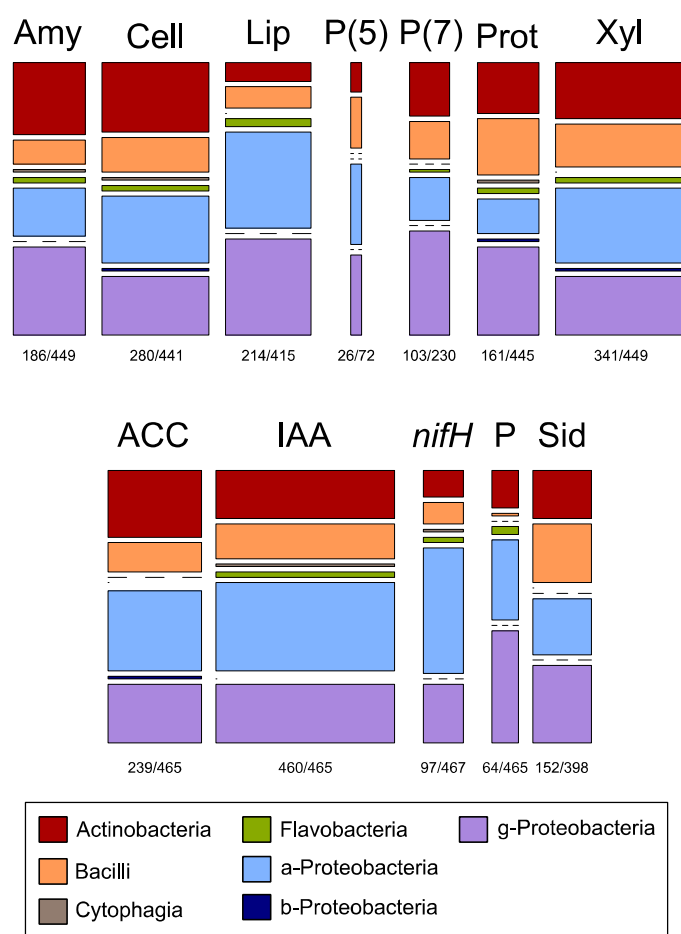
BBG-7	a-Proteobacteria	<i>Labrenzia</i>	4	KT324981	-	+	nd	+	+	nd	nd	-	6.5	-	+	+
BBG-8	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324982	nd	nd	nd	nd	+	nd	nd	-	10.2	-	-	+
BBG-9	Actinobacteria_c	<i>Demequina</i>	1	KT324983	-	+	-	+	+	nd	-	-	44.4	nd	-	-
BBG-10	g-Proteobacteria	<i>Shewanella</i>	2	KT324984	-	+	+	+	+	+	-	-	47.9	-	-	+
BBG-11	a-Proteobacteria	<i>Cohaesibacter</i>	1	KT324985	+	-	+	+	-	-	-	-	56.5	nd	+	-
BBG-12	a-Proteobacteria	<i>Aquamicrobium</i>	1	KT324986	-	-	-	+	+	nd	+	-	65.9	-	-	+
BBG-13	a-Proteobacteria	<i>Hoeflea</i>	4	KT324987	-	-	+	+	-	nd	+	-	92.3	-	-	-
BBG-14	a-Proteobacteria	<i>Sphingorabdus</i>	1	KT324988	-	-	+	+	+	nd	+	-	51.3	-	+	-
BBG-15	a-Proteobacteria	<i>Labrenzia</i>	2	KT324989	nd	nd	nd	nd	nd	nd	nd	-	57.6	-	-	+
BBG-16	g-Proteobacteria	<i>Vibrio</i>	1	KT324990	+	nd	+	+	+	+	+	-	4.7	nd	-	-
BBG-17	Actinobacteria_c	<i>Demequina</i>	1	KT324991	-	+	-	+	+	nd	-	-	42.1	-	-	-
BBG-18	a-Proteobacteria	<i>Labrenzia</i>	2	KT324992	nd	nd	nd	nd	nd	nd	-	-	100.4	+	+	+
BBG-19	a-Proteobacteria	<i>Labrenzia</i>	1	KT324993	-	-	-	+	+	nd	-	-	59.5	-	+	+
BBG-20	a-Proteobacteria	<i>Breoghania</i>	1	KT324994	-	-	+	+	+	-	+	-	65.4	-	+	+
BBG-21	nd	nd	1	nd	-	+	+	+	+	+	+	-	42.7	-	-	-
BBG-22	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT324995	-	+	+	-	+	-	+	-	23.2	-	-	-
BBG-23	a-Proteobacteria	<i>Labrenzia</i>	2	KT324996	-	+	-	+	+	nd	-	-	53.9	nd	+	+
BBG-24	Bacilli	<i>Bacillus</i>	1	KT324997	-	+	+	+	+	+	+	-	10.2	+	-	+
BBG-25	Bacilli	<i>Bacillus</i>	1	KT324998	-	+	+	+	+	+	+	-	10.2	+	-	+
BBG-26	Actinobacteria_c	<i>Demequina</i>	1	KT324999	-	+	-	+	+	nd	-	-	45.9	nd	-	-
BBG-27	a-Proteobacteria	<i>Rhizobium</i>	1	KT325000	-	-	+	+	+	-	-	+	62.4	+	-	+
BBG-28	Actinobacteria_c	<i>Demequina</i>	1	KT325001	-	+	-	+	+	-	nd	-	31.6	-	-	-
BBG-29	Actinobacteria_c	<i>Microbacterium</i>	1	KT325002	-	-	-	+	+	nd	-	+	29.7	+	-	+
BBG-30	Bacilli	<i>Bacillus</i>	1	KT325003	-	+	-	+	+	+	+	-	7.4	+	-	+
BBG-31	Actinobacteria_c	<i>Microbacterium</i>	3	KT325004	-	-	+	+	+	-	-	+	15.5	-	-	+
BBG-32	Bacilli	<i>Bacillus</i>	1	KT325005	-	+	+	+	+	nd	+	-	4.5	+	-	+
BBG-33	Bacilli	<i>Bacillus</i>	1	KT325006	-	+	+	+	+	nd	+	-	4.5	+	-	+
BBG-34	Actinobacteria_c	<i>Microbacterium</i>	3	KT325007	-	-	-	+	+	nd	-	-	31.9	+	-	+
BBG-35	g-Proteobacteria	<i>Vibrio</i>	1	KT325008	+	+	+	+	+	-	+	-	20.8	+	-	-
BBG-36	Bacilli	<i>Aerococcus</i>	2	KT325009	-	-	-	+	+	nd	+	-	4.9	-	-	+
BBG-37	a-Proteobacteria	<i>Hoeflea</i>	1	KT325021	-	-	-	+	+	nd	nd	-	65.5	+	-	-
BBG-38	a-Proteobacteria	<i>Martellella</i>	2	KT325022	-	-	-	-	-	nd	nd	-	24.2	-	-	-
BBG-39	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325023	-	-	+	-	+	nd	nd	-	9.7	-	-	+
BBG-40	Actinobacteria_c	<i>Demequina</i>	1	KT325024	-	-	-	+	+	nd	nd	-	55.2	nd	-	-
BBG-41	Actinobacteria_c	<i>Microbacterium</i>	2	KT325025	+	-	-	+	+	nd	+	-	1.3	+	-	+
BBG-42	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325026	+	-	+	+	nd	nd	-	-	6.6	+	-	-
BBG-43	a-Proteobacteria	<i>Rhizobium</i>	1	KT325027	-	-	-	+	+	nd	nd	+	54.8	-	-	-
BBG-44	a-Proteobacteria	<i>Citromicrobium</i>	1	KT325028	+	-	+	+	+	nd	nd	-	19.8	-	-	-
BBG-45	a-Proteobacteria	<i>Hoeflea</i>	2	KT325029	+	nd	+	-	+	nd	-	-	58.2	-	-	+

BBG-46	a-Proteobacteria	<i>Sulfitobacter</i>	1	KT325030	-	-	-	+	+	nd	nd	-	21.6	-	-	-
BBG-47	a-Proteobacteria	<i>Oceanicola</i>	1	KT325031	-	+	+	+	+	nd	nd	-	24.1	-	+	-
BBG-48	a-Proteobacteria	<i>Hoeflea</i>	1	KT325032	-	-	-	-	+	nd	nd	+	15.6	-	-	-
BBG-49	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325033	nd	+	+	+	+	nd	nd	-	23.9	-	-	+
BBG-50	a-Proteobacteria	<i>Aurantimonas</i>	1	KT325034	-	-	-	-	+	+	nd	+	27.1	-	+	+
BBG-51	a-Proteobacteria	<i>Labrenzia</i>	1	KT325035	-	-	-	+	+	nd	nd	-	57.9	nd	+	+
BBG-52	Flavobacteria	<i>Flavobacterium</i>	1	KT325036	+	+	+	+	+	nd	nd	-	48.3	nd	-	-
BBG-53	a-Proteobacteria	<i>Hoeflea</i>	2	KT325037	-	-	-	+	+	nd	nd	-	33.7	-	-	+
BBG-54	a-Proteobacteria	<i>Breoghania</i>	1	KT325038	-	-	-	+	+	nd	-	+	5.7	-	+	+
BBG-55	Bacilli	<i>Bacillus</i>	1	KT325039	+	+	+	+	+	nd	nd	-	39.2	+	-	+
BBG-56	a-Proteobacteria	<i>Martellella</i>	1	KT325040	-	-	-	-	-	-	-	-	25.6	-	-	-
BBG-57	Actinobacteria_c	<i>Demequina</i>	1	KT325041	-	-	-	+	+	nd	-	-	1.4	-	-	+
BBG-58	a-Proteobacteria	<i>Sulfitobacter</i>	1	KT325042	-	-	-	+	+	nd	nd	-	30.9	+	-	+
BBG-59	a-Proteobacteria	<i>Aquamicrobium</i>	1	KT325043	+	-	-	-	+	nd	nd	-	54.3	nd	+	+
BBG-60	Actinobacteria_c	<i>Microbacterium</i>	1	KT325044	-	-	-	+	+	nd	-	+	4.9	-	-	+
BBG-61	Actinobacteria_c	<i>Demequina</i>	2	KT325045	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	-	nd
BBG-62	Actinobacteria_c	<i>Microbacterium</i>	8	KT325046	-	-	-	+	+	nd	-	+	5.9	-	-	+
BBG-63	Bacilli	<i>Bacillus</i>	1	KT325047	-	+	-	+	+	+	+	-	9.3	+	-	+
BBG-64	a-Proteobacteria	<i>Altererythrobacter</i>	2	KT325048	-	+	+	+	+	+	+	-	8.3	+	-	+
BBG-65	a-Proteobacteria	<i>Altererythrobacter</i>	3	KT325049	+	-	+	+	+	nd	+	-	16.0	+	-	+
BBG-66	g-Proteobacteria	<i>Marinomonas</i>	1	KT325050	+	+	+	+	+	+	+	-	10.3	-	-	+
BBG-67	a-Proteobacteria	<i>Altererythrobacter</i>	3	KT325051	-	-	+	+	+	nd	-	-	7.4	+	-	+
BBG-68	g-Proteobacteria	<i>Pseudomonas</i>	2	KT325052	-	-	+	+	+	nd	nd	+	25.0	-	-	+
BBG-69	a-Proteobacteria	<i>Aurantimonas</i>	1	KT325053	-	-	-	-	+	+	nd	+	67.7	-	-	+
BBG-70	a-Proteobacteria	<i>Aurantimonas</i>	1	KT325054	-	-	-	-	-	-	nd	+	39.3	-	+	+
BBG-71	a-Proteobacteria	<i>Hoeflea</i>	1	KT325071	-	-	-	-	+	nd	nd	-	43.5	-	-	+
BBG-72	g-Proteobacteria	<i>Pseudomonas</i>	1	KT325072	+	-	+	-	+	-	nd	+	7.3	-	-	+
BBG-73	a-Proteobacteria	<i>Thalassospira</i>	2	KT325073	-	+	-	+	+	+	nd	-	25.4	+	-	-
BBG-74	a-Proteobacteria	<i>Rhodovulum</i>	2	KT325074	-	-	-	+	+	nd	nd	-	47.0	-	+	+
BBG-75	a-Proteobacteria	<i>Altererythrobacter</i>	1	KT325075	-	+	+	-	+	nd	nd	-	90.8	-	-	-
BBG-76	a-Proteobacteria	<i>Hoeflea</i>	1	KT325076	-	-	-	-	+	nd	nd	-	33.0	-	-	+
BBG-77	a-Proteobacteria	<i>Cohaesibacter</i>	1	KT325077	-	-	+	-	-	nd	nd	+	37.3	-	+	+
BBG-78	a-Proteobacteria	<i>Labrenzia</i>	1	KT325078	-	-	-	-	+	nd	nd	-	102.9	nd	+	+
BBG-79	Flavobacteria	<i>Muricauda</i>	1	KT325079	-	-	+	+	+	nd	nd	-	34.6	nd	-	-
BBG-80	Flavobacteria	<i>Muricauda</i>	1	nd	-	-	+	+	+	nd	nd	-	34.6	nd	-	-
BBG-81	Actinobacteria_c	<i>Demequina</i>	1	KT325080	+	-	-	+	+	nd	nd	-	37.3	nd	-	-
BBG-82	Flavobacteria	<i>Zunongwangia</i>	1	KT325081	nd	nd	nd	nd	nd	nd	nd	-	43.5	-	-	-
BBG-83	g-Proteobacteria	<i>Vibrio</i>	1	KT325082	+	+	+	+	+	nd	nd	-	7.1	+	-	-
BBG-84	g-Proteobacteria	<i>Vibrio</i>	4	KT325083	+	+	+	+	+	nd	-	-	3.5	-	-	+

BBG-85	g-Proteobacteria	<i>Vibrio</i>	1	KT325084	+	+	+	+	+	nd	-	-	11.7	-	+	-
BBG-86	g-Proteobacteria	<i>Vibrio</i>	1	KT325085	+	+	+	+	+	nd	-	-	42.7	-	+	-
BBG-87	g-Proteobacteria	<i>Vibrio</i>	1	KT325086	+	+	+	+	+	nd	nd	-	5.2	-	+	-
BBG-88	g-Proteobacteria	<i>Vibrio</i>	1	KT325087	+	-	+	+	+	nd	-	-	33.2	+	+	+
BBG-89	a-Proteobacteria	<i>Rhizobium</i>	2	KT325088	-	-	-	-	+	nd	nd	-	31.6	+	+	-
BBG-90	a-Proteobacteria	<i>Ensifer</i>	1	KT325089	+	-	+	+	+	-	-	+	108.5	+	+	+
BBG-91	a-Proteobacteria	<i>Stakelama</i>	1	KT325090	-	-	+	+	+	-	-	+	91.7	+	+	+
BBG-92	Actinobacteria_c	<i>Demequina</i>	1	KT325091	-	-	-	+	+	nd	+	-	42.8	-	+	-
BBG-93	g-Proteobacteria	<i>Pseudomonas</i>	2	KT325092	-	-	+	+	+	nd	+	+	32.6	-	-	+
BBG-94	Actinobacteria_c	<i>Microbacterium</i>	1	KT325093	-	-	-	+	+	nd	-	-	16.9	-	+	+
BBG-95	Bacilli	<i>Bacillus</i>	1	KT325094	+	+	-	+	+	nd	nd	-	19.3	+	+	+
BBG-96	g-Proteobacteria	<i>Vibrio</i>	1	KT325095	+	-	+	+	+	nd	-	-	21.6	-	+	+
BBG-97	g-Proteobacteria	<i>Halomonas</i>	1	KT325096	+	-	+	+	+	nd	+	-	16.2	+	+	+
BBG-98	a-Proteobacteria	<i>Altererythrobacter</i>	10	KT325097	-	-	+	+	+	-	-	+	17.8	+	+	+

CAG, aboveground (AG) tissues from site C; CBG, belowground (BG) tissues from site C; EAG, AG tissues from site E; EBG, BG tissues from site E; BAG, AG tissues from site B; BBG, BG tissues from site B; nd, not determined; a-, b- and g-Proteobacteria refers to Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria, respectively; IAA production is given in  $\mu\text{g mL}^{-1}$ .





**Supplementary Figure S3.1** Distribution of isolates that tested positive for enzymatic activity assays and plant growth promotion traits from 467 representative isolates. Isolates are grouped according to class-level taxonomic identification. Numbers below distributions refer to the number of representative isolates that tested positive/number of successfully tested isolates for that test. Amy, amylolytic activity; Cell, cellulolytic activity; Lip, lipolytic activity; P(5), pectinolytic at pH 5.0; P(7), pectinolytic at pH 7.0; Prot, proteolytic activity; Xyl, xylanolytic activity; ACC, growth in DF+ACC medium; IAA, IAA production; *nifH*, presence of *nifH* gene; P, phosphate solubilization; Sid, Siderophore production. a-, b- and g-Proteobacteria refers to Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria, respectively.

### **3.2 Diversity of endophytic *Pseudomonas* in *Halimione portulacoides* from metal(loid)-polluted salt marshes**

#### **Authors**

Jaqueline Rocha<sup>1</sup>, Marta Tação<sup>2</sup>, Cátia Fidalgo<sup>2</sup>, Artur Alves<sup>1</sup>, Isabel Henriques<sup>2</sup>

<sup>1</sup> Biology Department and CESAM, University of Aveiro, 3810-193 Aveiro, Portugal

<sup>2</sup> Biology Department, CESAM and iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal

#### **Publication status**

Published

Environ Sci Pollut Res

2016 Jul; 23(13):13255-67

doi: 10.1007/s11356-016-6483-x

## Abstract

Phytoremediation assisted by bacteria is seen as a promising alternative to reduce metal contamination in the environment. The main goal of this study was to characterize endophytic *Pseudomonas* isolated from *Halimione portulacoides*, a metal-accumulator plant, in salt marshes contaminated with metal(loid)s. Phylogenetic analysis based on 16S rRNA and *gyrB* genes showed that isolates affiliated with *P. sabulinigri* ( $n = 16$ ), *P. koreensis* ( $n = 10$ ), *P. simiae* ( $n = 5$ ), *P. seleniipraecipitans* ( $n = 2$ ), *P. guineae* ( $n = 2$ ), *P. migulae* ( $n = 1$ ), *P. fragi* ( $n = 1$ ), *P. xanthomarina* ( $n = 1$ ), and *Pseudomonas* sp. ( $n = 1$ ). Most of these species have never been described as endophytic. The majority of the isolates were resistant to three or more metal(loid)s. Antibiotic resistance was frequent among the isolates but most likely related to species-intrinsic features. Common acquired antibiotic resistance genes and integrons were not detected. Plasmids were detected in 43.6 % of the isolates. Isolates that affiliated with different species shared the same plasmid profile but attempts to transfer metal resistance to receptor strains were not successful. Phosphate solubilization and IAA production were the most prevalent plant growth promoting traits, and 20 % of the isolates showed activity against phytopathogenic bacteria. Most isolates produced four or more extracellular enzymes. Preliminary results showed that two selected isolates promote *Arabidopsis thaliana* root elongation. Results highlight the diversity of endophytic *Pseudomonas* in *H. portulacoides* from contaminated sites and their potential to assist phytoremediation by acting as plant growth promoters and as environmental detoxifiers.

## Keywords

*Pseudomonas*, *Halimione portulacoides*, endophytic, phytoremediation, plant growth promoters, metals

## Background

The contamination of different aquatic and terrestrial settings with metals is a complex and difficult to handle environmental problem (Cambrollé et al., 2012a; Rajkumar et al., 2009; Singh and Cameotra, 2004). Metal pollution is of major concern for human health due to the persistence of these elements in the environment for long periods, resulting in continuous negative impacts on ecosystems (Rajkumar et al., 2009; Singh and Cameotra, 2004).

Salt marshes are one of the most productive ecosystems on Earth. Nevertheless, these ecosystems are recipients of industrial and municipal waste, accumulating considerable amounts of contaminants, including metals (Duarte et al., 2007; Mucha et al., 2011; Sousa et al., 2008).

The strategies available to solve the problem of metal contamination include phytoremediation. This technique is considered effective, less expensive, and less damaging than, for example, chemical and physical remediation strategies (Almeida et al., 2009; Mucha et al., 2011; Weyens et al., 2009). Plants used in phytoremediation should have specific features, such as fast growth, and dense root and shoot systems. The phytoremediation of metal-contaminated sites is possible since some plants tolerate and accumulate high amounts of metals (Rajkumar et al., 2009). *Halimione portulacoides* is one of the most predominant and productive plants in European salt marshes and is already used in phytoremediation (Anjum et al., 2011; Carvalho et al., 2010; Couto et al., 2011). This plant has been reported as an accumulator of several metals and it has been proposed as biomonitor for mercury (Hg) contamination (Cambrollé et al., 2012a, b; Válega et al., 2008a).

The phytoremediation potential of a given plant also depends on the bacterial communities that live in association with it. Bacteria that are present in the phytosphere may contribute to phytoremediation through mechanisms that enhance plant establishment, proliferation, and health in a polluted environment. Beneficial effects provided by bacteria include increased metal tolerance, production of plant growth promoters (PGP), and antimicrobial activity against phytopathogens (Rajkumar et al., 2009). The contribution may be more direct since, for instance, some bacteria can modify the metals and thus facilitate absorption by the plant (Aguilar-Barajas et al., 2010; Ullah et al., 2015).

The genus *Pseudomonas* includes bacteria that can be found in a variety of ecological niches such as soil, water, and in association with plants and animals (Spiers et al., 2000). *Pseudomonas* spp. are a highly diverse and adaptable group, including human pathogens (e.g., *Pseudomonas aeruginosa*; Poole, 2011) and plant pathogens (e.g., *Pseudomonas syringae*; Tarkowski and

Vereecke, 2014). Association of *Pseudomonas* strains with plants has been referred as beneficial through their PGP effect (Ganeshan and Kumar, 2005). Despite this, endophytic *Pseudomonas* have been poorly explored regarding their potential to promote bioremediation when associated with plants.

The objective of this study was to characterize endophytic *Pseudomonas* isolates of *H. portulacoides* from salt marshes contaminated with metals. Isolates were comprehensively characterized with respect to their phylogenetic diversity, PGP effect, and resistance/tolerance to metals and antibiotics.

## Methods

### Endophytic *Pseudomonas* strains

A collection of endophytic bacterial isolates was obtained in the scope of previous investigations (Fidalgo et al., 2016; Martins, 2011). Bacterial isolates were retrieved from above (AG) and belowground (BG) tissues of *H. portulacoides* in metal-contaminated salt marshes (sites B and E) of the estuary Ria de Aveiro in the northwest coast of Portugal. Previous studies have shown that site B, being closest to the industrial effluent discharge point, showed higher metal(loid)s content than site E (Válega et al., 2008b). Bacterial isolation methodology is described in detail by Fidalgo et al. (2016). Briefly, samples were washed with tap water and for endophytic bacteria isolation, AG and BG tissues from each plant were used. Plant surface was sterilized by sequentially immersing samples in 50 mL of phosphate-buffered saline (PBS) for 10 min, 96 % ethanol during 1 min, 5 % NaOCl 30 min, 96 % ethanol 1 min, and rinsing in distilled water three times. Sterilization was confirmed by inoculating water from the last wash in adequate culture medium. Afterwards, tissues were then ground with a pestle in a mortar containing 10 mL of PBS; dilutions were prepared and spread on Tryptic Soy Agar (TSA, Merck, Germany), Marine Agar (MA, Difco, France), and R2A Agar (Merck, Germany) culture media. The plates were incubated up to 72 h at 28 °C and observed for colony forming units count every 24 h. Endophytic bacterial colonies observed as distinct morphologically were selected, purified, and stored at –80 °C in 20 % glycerol.

After molecular typing by BOX-PCR, representative isolates were identified at the genus level by sequencing about 800 bp of the 16S rRNA gene (Fidalgo et al., 2016; Martins, 2011). Isolates identified as *Pseudomonas* ( $n = 39$ ) were selected for further analysis in the scope of the present study.

### Phylogenetic analysis

Phylogenetic diversity within the collection of *Pseudomonas* isolates was assessed using nearly full-length 16S rRNA gene sequences. For this, the 16S rRNA gene was amplified with primers and PCR conditions described before (Tação et al., 2015). Similarity searches were performed with the BLAST software (Altschul et al., 1990) against the GenBank database. Closest relatives were also determined with the EZTaxon tool available at <http://www.ezbiocloud.net/eztaxon> (Kim et al., 2012). Additionally, *gyrB* gene of *Pseudomonas* isolates was sequenced, using PCR primers and conditions described elsewhere (Yamamoto et al., 2000).

Maximum-likelihood and neighbor-joining trees were constructed based on the concatenated sequences of 16S rRNA and *gyrB* genes of isolates. Sequences of the corresponding genes of closest relative type strains of *Pseudomonas* spp. (according to BLAST and EzTaxon results) were used to construct phylogenetic trees using molecular evolutionary genetics analysis version 6.0 (MEGA6; Tamura et al., 2013). Kimura-2-parameter was used as the model of DNA sequence evolution. The branch numbers refer to bootstrap confidence values obtained using 1000 replications.

### **Metal tolerance assays**

Minimal inhibitory concentrations (MICs) were determined in Luria-Bertani (LB, Fisher BioReagents, Belgium) agar supplemented with arsenic (As; 50, 100, 300, and 600  $\mu\text{g mL}^{-1}$  As as  $\text{Na}_2\text{AsO}_4$ ), chromium (Cr III; 50, 100, 300, and 600  $\mu\text{g mL}^{-1}$  Cr as  $\text{CrCl}_3 \cdot 6\text{H}_2\text{O}$ ), copper (Cu; 100, 300, 400, 600, and 1000  $\mu\text{g mL}^{-1}$  Cu as  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ), mercury (Hg; 5, 100, 150, 250, and 400  $\mu\text{g mL}^{-1}$  Hg as  $\text{HgCl}_2$ ), nickel (Ni; 50, 100, 300, and 600  $\mu\text{g mL}^{-1}$  Ni as  $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ ), and zinc (Zn; 50, 100, 300, and 600  $\mu\text{g mL}^{-1}$  Zn as  $\text{ZnCl}_2$ ). Metal(lloid) stock solutions were prepared in distilled water and sterilized. Experiments were conducted in triplicate. Results were registered after 5 days incubation at 30 °C. The reference strain *Escherichia coli* ATCC 25922 was included for quality control. Since there are no standard interpretative criteria for the classification of bacterial isolates as susceptible or resistant to metals, thresholds were defined based on previous studies (Nies, 1999; Malik and Aleem, 2011). Hence, the sensitivity thresholds considered were 600  $\mu\text{g mL}^{-1}$  for As, 20  $\mu\text{g mL}^{-1}$  for Hg, 250  $\mu\text{g mL}^{-1}$  for Cr, and 300  $\mu\text{g mL}^{-1}$  for Cu, Ni, and Zn.

### **Antibiotic susceptibility assays**

Isolates were tested for susceptibility to 16 antibiotics using the disk diffusion method on Mueller-Hinton Agar (Oxoid, Basingstoke, UK), following the Clinical Laboratory Standards Institute (CLSI) guidelines (CLSI 2012). The tested antibiotics were cefepime (FEP, 30  $\mu\text{g}$ ), ampicillin (AMP, 10  $\mu\text{g}$ ), amoxicillin (AML, 10  $\mu\text{g}$ ), amoxicillin/clavulanic acid (AMC, 20  $\mu\text{g}/10 \mu\text{g}$ ), piperacillin (PRL, 100  $\mu\text{g}$ ), aztreonam (ATM, 30  $\mu\text{g}$ ), piperacillin/tazobactam (TZP, 100  $\mu\text{g}/10 \mu\text{g}$ ), cefotaxime (CTX, 30  $\mu\text{g}$ ), ceftazidime (CAZ, 30  $\mu\text{g}$ ), imipenem (IPM, 10  $\mu\text{g}$ ); tetracycline (TE, 30  $\mu\text{g}$ ), gentamicin (CN, 10  $\mu\text{g}$ ), tobramycin (TOB, 10  $\mu\text{g}$ ), ciprofloxacin (CIP, 5  $\mu\text{g}$ ), chloramphenicol (C, 30  $\mu\text{g}$ ), and sulfamethoxazole/trimethoprim (SXT, 25  $\mu\text{g}$ ). After 24 h of incubation at 37 °C, organisms were classified as sensitive, intermediate, or resistant according to the CLSI guidelines. The *E. coli* ATCC

25922 was used as quality control. Cluster analysis of the antibiotic susceptibility profiles of *Pseudomonas* isolates was performed using Bray-Curtis similarity coefficient and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) algorithm. For this, matrices were constructed considering isolates classification as resistant (2), intermediate resistant (1), or susceptible (0) to antibiotics.

### **Plant growth promoting trait assays**

All isolates were screened in triplicate for different PGP activities: 1-aminocyclopropane-1-carboxylate (ACC) deaminase, phosphate solubilization, indole acetic acid (IAA), siderophore production, nitrogen fixation, and hydrogen cyanide (HCN) production. ACC deaminase activity was determined by growing isolates in DF salt minimal medium (Dworkin and Foster, 1958) supplemented with 3 mM ACC (DF+ACC) or 2 g L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (DF+Ammonium Sulfate) or no nitrogen source (DF). The ability to use ACC as a sole nitrogen source is a consequence of the activity of the enzyme ACC deaminase (Penrose and Glick, 2003). Phosphate solubilization was observed as a clear halo around bacterial growth on National Botanical Research Institute's Phosphate (NBRIP) solid growth medium (Nautiyal, 1999). IAA production was tested by the method described by Gordon and Weber (1951). Siderophore production was evaluated after inoculation in TSA medium followed by an overlay with O-CAS (Pérez-Miranda et al., 2007; Pérez-Miranda and Fernández, 2013) after incubation for 16 h at 30 °C. Bacterial growth overlaid with CAS was incubated for 2 h at room temperature. Siderophore production was identified as a change of color of the overlay from blue to orange or purple. Nitrogen fixing capacity was evaluated by PCR, targeting *nifH* gene with primers and conditions described elsewhere (Ando et al., 2005). HCN production was evaluated streaking the isolates on *Pseudomonas* agar base (Liofilchem, Italy) supplemented with 4.4 g L<sup>-1</sup> glycine. A sterile filter paper was soaked with a picric acid solution (2.5 g L<sup>-1</sup> picric acid and 12.5 g L<sup>-1</sup> Na<sub>2</sub>CO<sub>3</sub> on distilled water) and placed on the upper lid of each plate. The plates were sealed with parafilm and incubated during 48 h at 30 °C (Reetha et al., 2014). HCN production was qualitatively identified as a change of color of the filter paper from yellow to light brown, brown, and reddish brown. *Pseudomonas putida* EAPC8 and *Pseudomonas fluorescens* S3X were used as positive controls for IAA production. For phosphate solubilization "*Pseudomonas reactans*" EDP28, *P. putida* EAPC8, *Arthrobacter rombi* EC32A, and *P. fluorescens* S3X were used as negative controls, and *Arthrobacter nicotinovorans* EAPAA and *Rhodococcus* sp. EC35 were used as positive controls. For siderophore production "*P. reactans*" EDP28, *P. putida* EAPC8, *A. rombi* EC32A, *P. fluorescens*



S3X, *A. nicotinovorans* EAPAA, and *Rhodococcus* sp. EC35 were used as positive controls (Pereira et al., 2015). *Rahnella aquatilis* (HQ538817; Proença et al., 2010) was used as a positive control and *E. coli* ATCC 25922 was used as a negative control (Gaby and Buckley, 2012). *Pseudomonas putida* EAPC8 and *A. rombi* EC32A were used as positive controls for the ACC deaminase activity tests. For HCN production, a plate without bacterial suspension was used as negative control.

### **Antimicrobial activity assays**

Antimicrobial activity was tested against six indicator strains from genera that include phytopathogenic members, isolated from *H. portulacoides* in the scope of previous studies (Martins, 2011) or obtained from private collections: *Wautersia* (*W. eutropha* NR92), *Agrobacterium* (*A. tumefaciens* NR163SFP) *Pantoea* (*Pantoea brenneri* PA112, *Pantoea eucalypti* RZ12) *Erwinia* (*Erwinia rhapontici* RZ93), and *Curtobacterium* (*Curtobacterium flaccumfaciens* PA10). Assays were performed in TSA medium. The indicator strain was inoculated in TSA plates and left at room temperature for 30 min. These plates were then inoculated with 5 µL of a bacterial suspension of the collection of endophytic bacteria in NaCl 0.9 % (1 McFarland) and incubated at 30 °C for 16 h. Antimicrobial activity was observed according to the presence (positive result) or absence (negative result) of a halo surrounding the bacterial growth (Prado et al., 2009).

### **Enzymatic assays**

All isolates were screened, in triplicate, for the following enzymatic activities: amylolytic, cellulolytic, lipolytic, pectinolytic, proteolytic, and xylanolytic. For enzymatic activity screening, bacterial isolates were streaked on TSA culture medium supplemented with appropriate substrate: 0.2 % starch for amylolytic activity, 0.5 % carboxymethyl cellulose for cellulolytic activity, 1 % Tween 20 for lipolytic, 0.5 % pectin for pectinolytic, 1 % skim milk for proteolytic, and 0.5 % xylan for xylanolytic activity. Amylolytic activity was detected when a clear halo was observed upon addition of Lugol solution after bacterial growth. Cellulolytic and xylanolytic activities were detected when a clear halo was observed upon incubation in TSA with Congo red solution (1 % Congo red, with 10 % ethanol) and subsequent wash with 1 M NaCl. Lipolytic activity was observed by the formation of a white precipitate around the bacterial growth. Pectinolytic activity was detected by formation of a clear halo after addition of 1 % hexadecyltrimethylammonium bromide (CTAB) solution. Proteolytic activity was revealed by visualization of a clear halo against an opaque background. The

results were interpreted according to the visibility of the halo obtained: negative, no halo observed; positive, faint to intense halo observed.

### **Effect of bacteria inoculation on *Arabidopsis thaliana* root elongation**

The effect of bacterial inoculation on *A. thaliana* root growth was evaluated using a protocol based on Shi et al. (2010). The disinfection of the *A. thaliana* seeds was processed in sterile microtubes with 500 µL of a mixture of 70 % ethanol and 0.05 % Tween 20, followed by 5 min incubation at room temperature. The supernatant was aspirated with a micropipette and 500 µL of 100 % ethanol were added. After an incubation of 5 min at room temperature, the ethanol was discarded. The disinfected seeds ( $n = 15$ ) were spread in a single line on half strength Murashige and Skoog medium (MS, Duchefa Biochemie, The Netherlands), supplemented with 1.5 % glucose and 0.8 % agar and the plates were incubated, vertically, for 3 days at 4 °C, to promote germination. A bacterial cell suspension on LB broth ( $OD_{600} = 0.6–0.8$ ) was pelleted during 1 min at 13000 rpm and resuspended in 1 mL of half strength MS medium supplemented with 1.5 % glucose and without agar. Sterile filter papers with 1 cm diameter were placed at 2 cm distance from the seeds line with 50 µL of bacterial suspension. Plates were incubated, vertically, at 24 °C under a 16/8 h light-dark cycle. The tests were performed in triplicate and a negative control without bacterial suspension was included. At the end of the incubation period, 15 roots were measured and statistical analysis was performed. The normality of the distribution of the root measurements was assessed with Shapiro's normality test (*shapiro.test*). If a normal distribution was observed then the significance of the differences for each treatment against the control was assessed with *t* student test (*t.test*). Tests were performed using the *stats* package in R (R Core Team, 2014).

### **PCR screening for integrons, metal, and antibiotic resistance genes**

Genes conferring resistance to beta-lactams (*bla<sub>SHV</sub>*, *bla<sub>TEM</sub>*, *bla<sub>CTX-M</sub>*, *bla<sub>GES</sub>*, *bla<sub>KPC</sub>*, *bla<sub>VIM</sub>*, *bla<sub>IMP</sub>*, *bla<sub>OXA-2-like</sub>*, *bla<sub>OXA-10-like</sub>*, *bla<sub>OXA-30-like</sub>*), sulphonamides (*sul1*), and chloramphenicol (*cat*) were inspected by PCR. Integrase screening was performed for *int1* and *int2* genes. All experiments included positive controls and also a PCR mixture containing water instead of DNA as negative control. Primers and PCR conditions are listed in Supplementary Table S3.4.

Mercury resistance gene *merA* (encoding a mercuric reductase) was inspected by PCR. Primers and conditions were used as previously described (Deredjian et al., 2011).

### **Plasmid profiles and replicon typing**

Bacterial plasmid DNA was extracted with E.Z.N.A. Plasmid Mini Kit II (Omega Bio-tek, GA, USA). Plasmid diversity was analyzed by restriction analysis using 5 U of *Pst*I (CTGCA↓G) and 5 U of *Bst*1770I (GTA↓TAC), according to the manufacturer's instructions (Fermentas, Lithuania). Restriction patterns were visualized in 0.8 % agarose gels (electrophoresis run at 40 V for 3 h) stained with ethidium bromide, and visualized with the Molecular Imager® Gel Doc™ XR System and Image Lab™ Software (Bio-Rad, Hercules, CA, USA). Restriction patterns were matched using GelCompar II software (Applied Maths, Sint-Martens-Latem, Belgium).

Replicon typing was performed by PCR for detection of FIA, FIB, FIC, HI1, HI2, I1-Ig, L/M, N, P, W, T, A/C, K, B/O, X, Y, F, and FIIA replicons, using primers and conditions described by Carattoli et al. (2005).

Conjugation experiments were performed for selected isolates that showed resistance to Hg. This metal was chosen since study sites were subjected to Hg-rich industrial discharges for four decades (1950s–1990s, Costa and Jesus-Rydin, 2001). For that, log-phase cultures of donor isolates and the recipient strain *E. coli* CV601 were mixed at a 1:1 ratio in LB (Merck, Germany) broth and incubated at 30 °C for 24 h without shaking. Transconjugant selection was performed with LB agar plates with 50 µg mL<sup>-1</sup> rifampicin and 25 µg mL<sup>-1</sup> Hg as HgCl<sub>2</sub>.

### **Nucleotide sequence accession numbers**

All nucleotide sequences reported in this work have been submitted to the GenBank database, under the accession numbers KT710788-KT710865.

## Results and Discussion

The environmental contamination with metals is a worldwide problem. The use of plants to remove metals from polluted environments, i.e., phytoremediation, is seen as a promising alternative. Endophytic bacteria, which colonize internal tissues of plants without causing damage to the host, have been reported as facilitators of phytoremediation processes (Compant et al., 2010), by promoting the proliferation of plants in contaminated sites.

The genus *Pseudomonas* is a diverse microbial group that, due to its high genome plasticity, is able to adapt to severe environments and to degrade or tolerate a range of natural and synthetic compounds (Aguilar-Barajas et al., 2010). Additionally, members of this genus were reported to efficiently promote plant growth (Preston, 2004). In this study, we analyzed endophytic *Pseudomonas* isolates associated to the salt marsh plant *H. portulacoides* from metal(loid)-polluted sites.

Though the potential of this plant for remediation of metal-polluted sites has been recognized, its phytosphere has never been explored to identify bacterial strains that may contribute to this decontamination. Also, the specificities of salt marshes (e.g., high salinity, tide variation) justify the need to identify indigenous bacterial strains that can be used for phytoremediation in this particular environment.

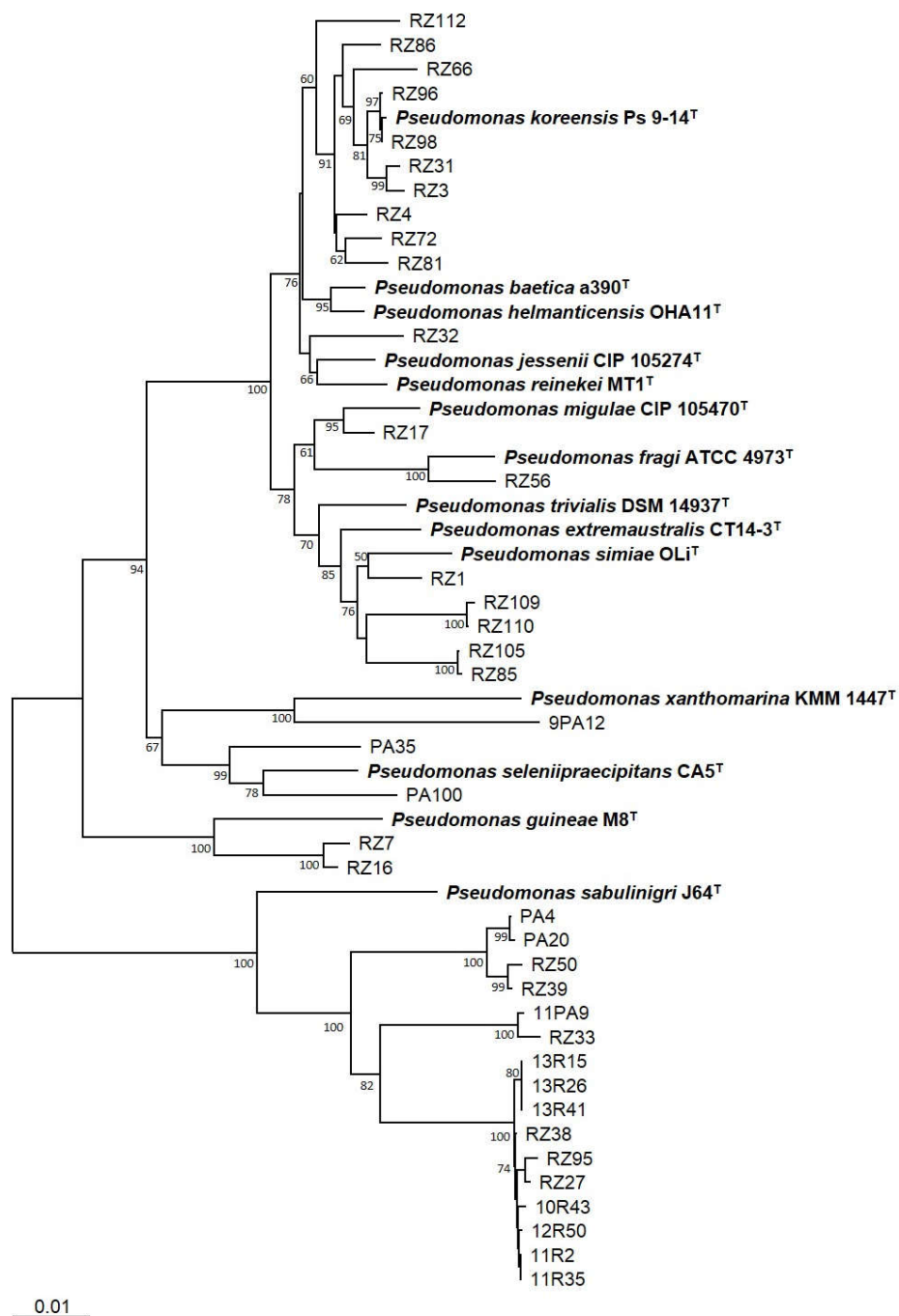
### Phylogenetic diversity of endophytic *Pseudomonas* isolates

A collection of 39 previously obtained endophytic isolates of *Pseudomonas* (Fidalgo et al., 2016; Martins, 2011) was used in this study. While in the study conducted by Martins et al. (samples collected in 2010), *Pseudomonas* was the predominant genus (46 % of the total number of isolates); in the study conducted by Fidalgo et al. (samples collected in 2012), *Pseudomonas* isolates represented only 2 % of the total number of isolates. The most common genera in the last study were *Microbacterium* and *Salinicola* representing 10 and 9 % of the collection, respectively. Differences in *Pseudomonas* prevalence between studies are probably related with the use of different culture media. While TSA was used in both studies, Fidalgo et al. used also Marine Agar and R2A Agar. Marine Agar presents a high salt content thus favoring isolation and cultivation of marine bacteria. R2A agar is a low nutrient medium specifically designed to cultivate oligotrophic bacteria. On the contrary, TSA may select for fast-growing heterotrophic bacteria including many *Pseudomonas* strains.

Detailed phylogenetic analysis of the *Pseudomonas* isolates used in the present study ( $n = 39$ ) is presented in Figure 3.4. The phylogenetic tree based on neighbor-joining method was constructed using concatenated 16S rRNA and *gyrB* gene sequences from all isolates and from closest related *Pseudomonas* spp. type strains. Affiliation of the isolates was based on well-supported monospecific clades (bootstrap values above 50 %) in the obtained phylogenetic tree. Based on this criterion, isolates affiliated with the closest related species *Pseudomonas sabulinigri* ( $n = 16$ ), *Pseudomonas koreensis* ( $n = 10$ ), *Pseudomonas simiae* ( $n = 5$ ), *Pseudomonas seleniipraecipitans* ( $n = 2$ ), *Pseudomonas guineae* ( $n = 2$ ), *Pseudomonas migulae* ( $n = 1$ ), *Pseudomonas fragi* ( $n = 1$ ), and *Pseudomonas xanthomarina* ( $n = 1$ ). One isolate (RZ32) grouped with sequences from two different species in the same clade. Phylogenetic trees constructed using a different method, maximum-likelihood, showed identical affiliation (data not shown). Although the applied methodology allowed affiliating most isolates to a closest related species, further analysis must be conducted to confirm identification at species level. Multigene sequence analyses of 16S rRNA, ITS, *gyrB*, and *rpoD* have proven useful for *Pseudomonas* species identification. Furthermore, the basis to describe new bacterial species has been DNA-DNA hybridization. However, the fact that many isolates grouped in distinct clusters, distant to the closest type strain, suggests that these isolates probably represent undescribed *Pseudomonas* species.

Results suggest a high diversity of endophytic *Pseudomonas* spp. associated to the plant *H. portulacoides*. Although plant association is widely recognized among pseudomonads (Preston, 2004), studies reporting endophytic *Pseudomonas* spp. are rare (Ryan et al., 2008; Rajkumar et al., 2009). In fact, the majority of the species retrieved during this study have not been described before as endophytic. Exceptions were *P. koreensis* (Babu et al., 2015) and *P. migulae* (Ali et al., 2014). *Pseudomonas koreensis* has also been described as epiphytic (Lopez-Velasco et al., 2012). *Pseudomonas koreensis* and *P. migulae* were previously isolated from *Miscanthus sinensis* rhizosphere in metal(loid)-contaminated sites (Babu et al., 2015; Chaturvedi et al., 2006).

Clearly, isolates affiliating with *P. koreensis* and *P. sabulinigri* were more prevalent among our collection. *Pseudomonas sabulinigri* has previously been isolated from black beach sand (Kim et al., 2009), while *P. koreensis* has been isolated from multiple sources, and includes biosurfactant-producing strains (Babu et al., 2015; Lopez-Velasco et al., 2012; Toribio et al., 2011).



**Figure 3.4** Neighbor-joining (bootstrap=1000) tree illustrating the phylogenetic position of the *Pseudomonas* isolates obtained in this study and the closest related *Pseudomonas* type strains. The tree was constructed based on 16S rRNA and *gyrB* concatenated gene sequences.

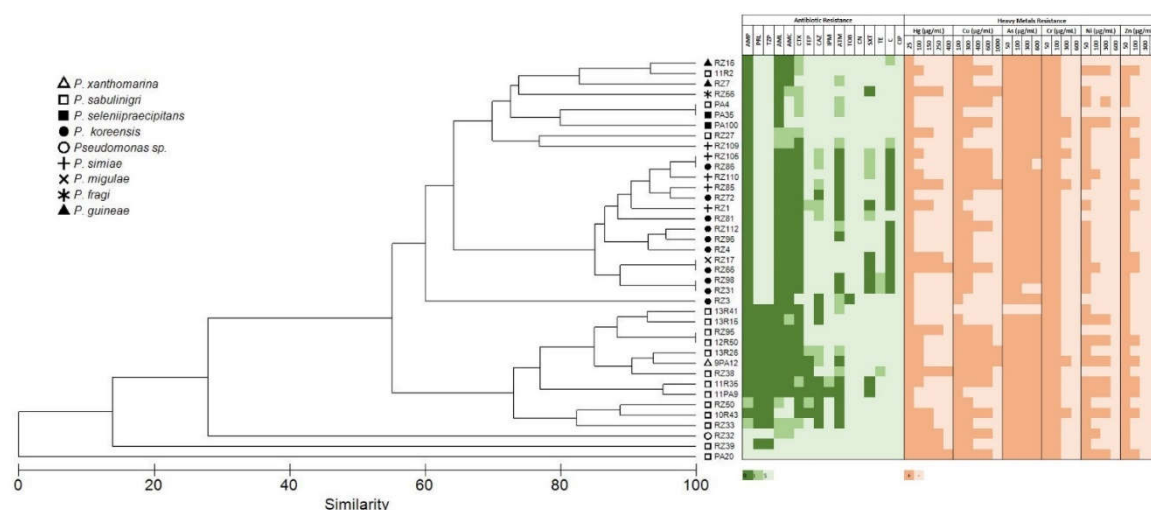
### **Metal tolerance, susceptibility to antibiotics, and plasmid profiling**

Endophytic bacteria may play an important role in enhancing the capability of plants to extract metals from polluted environments (Rajkumar et al., 2009). Tolerance to metal(loid)s (Hg, Ni, Cu, Cr, As, Zn) was studied in *Pseudomonas* isolates. The metal(loid) concentrations tested were higher than the maximum contamination limits suggested by the European Regulatory Standards (EURS) for European soils (Anyakora et al., 2013; Wuana and Okieimen, 2011). Additionally, these concentrations were higher than those reported for soil and plant tissues in the sampled sites (Fidalgo et al., 2016).

Bacterial tolerance to Hg, Cu, and As was high, with 100, 92, and 82 % of tolerant isolates, respectively (Figure 3.5). A lower percentage of tolerant isolates was observed for Ni (26 %) and for Cr (15 %). Tolerance to Zn was not observed. Overall, the order of metal(loid) tolerance among the isolates was Hg>As>Cu>Ni>Cr>Zn. The majority (82 %) of the isolates were tolerant to three or more metal(loid)s. Three isolates (PA100 closely related to *P. seleniipraecipitans*, 9PA12 closely related to *P. xanthomarina*, and 10R43 closely related to *P. sabulinigri*) showed tolerance to five metal(loid)s.

*Pseudomonas* have developed tolerance mechanisms against several metal(loid)s including micronutrient cations (Cu, Ni, and Zn), toxic cations (Hg), and toxic oxyanions (As and Cr) (Aguilar-Barajas et al., 2010; Stout and Nüsslein, 2010). A comparative genomics study (Wu et al., 2011), which included endophytic and rhizospheric *P. putida* strains, revealed that the endophytic strain presented the highest number of genes related to metal(loid) tolerance. Most of these genes were located in genomic islands. Thus, the endophytic strain seemed best equipped for dealing with metal(loid)-related pressure than its rhizospheric relative. In accordance, the endophytic *P. putida* strain tolerated the highest concentrations of metal(loid)s (Wu et al., 2011).

Isolates were most generally tolerant to Hg. One of the most common tolerance mechanisms to Hg is the reduction of this metal from the toxic cationic form to the volatile form. This biotransformation was described in several *Pseudomonas* species such as *P. fluorescens*, *P. putida*, and *Pseudomonas stutzeri* and is mediated by an Hg reductase encoded by *merA* gene located in the *mer* operon (Aguilar-Barajas et al., 2010). Even though we conducted a PCR screening for the *merA* gene, no positive results were obtained. It is plausible that the Hg-tolerance phenotypes here observed may be related to other mechanisms such as efflux systems, also commonly reported in *Pseudomonas* (Aguilar-Barajas et al., 2010).



**Figure 3.5** Cluster analysis of the antibiotic susceptibility profiles of *Pseudomonas* isolates from *H. portulacoides*, using Bray-Curtis similarity coefficient and unweighted pair group method, using arithmetic averages cluster methods. Results of metal(loid) tolerance assays are also presented. AMP, ampicillin; PRL, piperacillin; TZP, piperacillin/tazobactam; AML, amoxicillin; AMC, amoxicillin/clavulanic acid; CTX, cefotaxime; FEP, cefepime; CAZ, ceftazidime; IMP, imipenem; ATM, aztreonam; TOB, tobramycin; CN, gentamicin; CIP, ciprofloxacin; TE, tetracycline; C, chloramphenicol; SXT, sulfamethoxazole/trimethoprim; Hg, mercury; Cu, copper; As, arsenic; Ni, nickel; Zn, zinc; Cr, chromium.

*Pseudomonas* spp. tolerance to metals is extremely variable (Xu et al., 2015; Malik and Jaiswal, 2000; Shamim and Rehman, 2013). Endophytic *P. koreensis* isolates were reported to tolerate concentrations of Zn and As 30 and 4.5 times higher, respectively, than concentrations tolerated by our isolates (Babu et al., 2015). Those isolates were obtained from soil with concentrations of Zn and As 36 and 39 times higher than those reported for our sampling sites (Babu et al., 2015; Fidalgo et al., 2016). These results suggest that *Pseudomonas* community may adapt, to a certain limit, to deal with the metal(loid) pressure present in each environment.

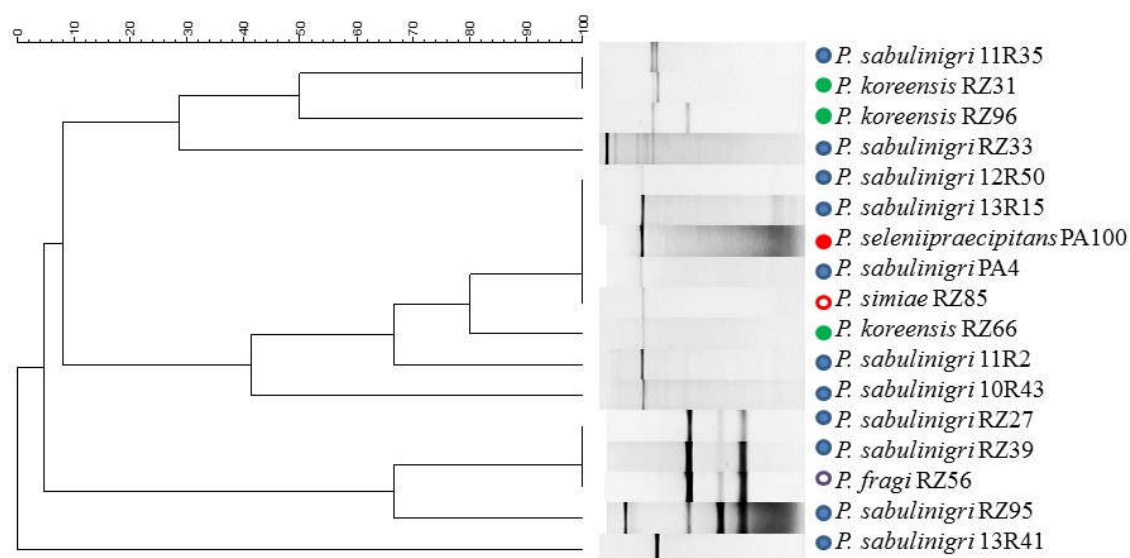
Tolerance to metal(loid)s may be accompanied by antibiotic resistance mechanisms that can be combined physiologically (cross-resistance) or genetically (co-resistance; Baker-Austin et al., 2006). Thus, the use of metal(loid)-tolerant *Pseudomonas* strains in phytoremediation may contribute to the spread of antibiotic resistance. From our collection, antibiotic susceptibility tests (Figure 3.5) showed higher resistance levels against penicillins, namely ampicillin and amoxicillin, to which 87 and 79 % of the isolates, respectively, were resistant. For non-beta-lactam antibiotics, higher resistance was observed to chloramphenicol (33 %) and sulfamethoxazole/trimethoprim (18 %).



Clustering analysis of all antibiotic resistance (Figure 3.5) resulted in a dendrogram where three main clusters were identified, corresponding to (i) isolates closely related to *P. sabulinigri*, *P. guineae*, *P. fragi*, *P. seleniipraecipitans*, and *P. simiae*, (ii) isolates closely related to *P. koreensis*, *P. simiae*, and *P. migulae*, and (iii) isolates closely related to *P. sabulinigri*, *P. xanthomarina*, and *Pseudomonas* sp.. Although a few exceptions could be identified in the dendrogram (e.g., not all isolates affiliated with *P. sabulinigri* grouped in the same cluster and a similar result was observed for *P. simiae*), the distribution of antibiotic resistance profiles was mostly consistent with the phylogenetic affiliation, suggesting that the observed antibiotic resistance is part of the intrinsic characteristics of each species. This hypothesis is corroborated by the fact that common acquired antibiotic resistance genes were not detected in the analyzed isolates.

Antibiotic multiresistance was observed for 13 % of the isolates, showing an intermediate or resistant phenotype at least to three classes of antibiotics. Regarding resistance mechanisms to antibiotics, *P. aeruginosa* is the most studied species in the genus *Pseudomonas*, due to its clinical importance. Multiresistance has been commonly reported in this species (Mesaros et al., 2007; Tacão et al., 2015) resulting from combinations of two or more resistance mechanisms (Livermore, 2001; Mesaros et al., 2007). These include overexpression of efflux pumps with wide substrate specificity, reduced outer membrane permeability, and/or high-level expression of  $\beta$ -lactamases (Livermore, 2001; Mesaros et al., 2007).

Acquired antibiotic resistance and metal(loid) resistance may be encoded in mobile genetic elements such as integrons and/or plasmids (Tacão et al., 2014). Integrase screening performed for *int1* and *int2* genes gave no positive results. Plasmidic DNA was identified in 17 out of 39 isolates (43.6 %) and enzymatic restriction analysis resulted in 10 different profiles. Plasmidic DNA restriction profiles were not species-specific since the same pattern (100 % identity) was obtained from isolates affiliated to different species (Figure 3.6). One banding profile was identified in both 11R35 (closely related to *P. sabulinigri*) and RZ31 (closely related to *P. koreensis*). Additionally, an identical restriction pattern was observed for two isolates closely related to *P. sabulinigri* (12R50 and PA4), RZ85 (affiliated with *P. simiae*), and PA100 (affiliated with *P. seleniipraecipitans*). Finally, identical profile was observed for two isolates closely related to *P. sabulinigri* (RZ27 and RZ39) and RZ56 (affiliated with *P. fragi*). These results suggest that different endophytic *Pseudomonas* species may share similar replicons. Conjugation assays were performed for Hg-resistant isolates but did not result in transconjugants and thus it was not possible to confirm the conjugative nature of such replicons.



**Figure 3.6** Cluster analysis of plasmidic DNA restriction profiles from endophytic *Pseudomonas* isolates. The Dice similarity coefficient was used, along with the unweighted pair group method, using arithmetic averages cluster methods.

### Plant growth promoting activity

In addition to conferring plants higher tolerance to metal (loid)s, endophytic bacteria play a key role in promoting plant growth through the acquisition of nutrients by nitrogen fixation and phosphate solubilization, and through the production of growth regulators such as IAA, siderophores, ACC deaminase, and HCN production (Chauhan et al., 2015; Long et al., 2008; Ma et al., 2011). The presence of PGP traits was assessed for all isolates. Each isolate displayed at least one PGP trait (Table 3.4). Prevalent PGP traits were IAA production (100 % of the isolates were positive for this trait), phosphate solubilization (90 % of the isolates), and HCN production (59 % of the isolates). Isolates producing concentrations of IAA higher than  $100 \mu\text{g mL}^{-1}$  were considered good producers, and represented 44 % of the collection. A lower fraction of isolates (44 %) was able to produce siderophores. PGP traits such as ACC-deaminase production (3 %) and the presence of *nifH* gene (5 %) were rarer in our collection. PGP traits were not species-specific, as can be observed in Table 3.4.

IAA is a phytohormone that increases plants biomass, thus contributing to increased metal accumulation in plants (Stout and Nüsslein, 2010). Bacterial fixation of atmospheric nitrogen and phosphate solubilization help plants in nutrient acquisition (Long et al., 2008). HCN promotes plant

growth preventing the damaging effect of phytopathogenic organisms (Chauhan et al., 2015). ACC deaminase lowers plant ethylene levels and improves plant growth in metal polluted conditions (Stout and Nüsslein, 2010). Siderophores, as chelating agents, protect plants, enabling them to resist to the hazardous effects of metals (Stout and Nüsslein, 2010).

In previous studies, the most common PGP traits detected in *Pseudomonas* spp. were ACC deaminase activity, siderophores, and IAA production, which are mainly involved in plant root and shoot elongation and metal uptake (Ma et al., 2011). Several *Pseudomonas* spp. have also been described as good phosphate solubilizers (Rajkumar et al., 2009) and HCN production is also common among pseudomonads (Pereira et al., 2016). In the present study, we observed that the majority of the isolates produced IAA, HCN, siderophores, and solubilized phosphate. This suggests that *Pseudomonas* strains in the endosphere may help plants accumulate metals, since plant biomass is increased, and metal toxicity to the plant is attenuated.

### **Enzymatic and antimicrobial activity**

Metal-tolerant bacterial strains have been reported as biological control agents, presenting antimicrobial activity against pathogenic fungi, bacteria, insects, and other organisms (Compant et al., 2010; Pereira et al., 2015).

Here, we evaluated antimicrobial activity against representatives of phytopathogenic bacteria. Antagonistic activity was observed in 20 % of the isolates, as shown in Table 3.4. RZ39 and RZ38 (closely related to *P. sabulinigri*), Z4 and RZ98 (affiliated with *P. koreensis*), and *Pseudomonas* sp. RZ32 showed antimicrobial activity against one tested strain. RZ31 (affiliated with *P. koreensis*) and RZ109 (affiliated with *P. simiae*) showed antimicrobial activity against two tested strains, and RZ1 (affiliated with *P. simiae*) against three test strains. These isolates may have a relevant role as biological control agents by protecting the plant against phytopathogens. So far, the evaluation of plant-associated pseudomonads as biocontrol agents has been mostly focused on *P. fluorescens* isolates, which have been proven to play an important role in antifungal activity (Petatán-Sagahón et al., 2011; Ganeshan and Kumar, 2005).

**Table 3.4** Plant growth promotion traits, antimicrobial, and enzymatic activities of *Pseudomonas* isolates from *H. portulacoides*.

Strains	Site	Plant Growth Promotion Traits						Antimicrobial Activity						Enzymatic Activity					
		P	IAA	Sid.	ACC	nifH	HCN	R. e.	A. t.	P. b.	P. e.	E. r.	C. f.	Prot	Lip	Cell	Pect	Amyl	Xyl
<i>P. sabulinigri</i> 10R43	B	+	+	-	-	-	-	-	-	-	-	-	-	-	+	+	nd	-	+
<i>P. sabulinigri</i> 11PA9	B	+	+	-	-	-	++	-	-	-	-	-	-	-	+	-	+	-	+
<i>P. sabulinigri</i> 11R2	B	+	+	-	-	-	+	-	-	-	-	-	-	-	+	-	nd	-	+
<i>P. sabulinigri</i> 11R35	B	+	+	-	-	+	+	-	-	-	-	-	-	-	+	+	+	-	+
<i>P. sabulinigri</i> 12R50	E	+	++	-	-	-	++	-	-	-	-	-	-	+	+	+	nd	-	+
<i>P. sabulinigri</i> 13R15	E	+	++	-	-	+	++	-	-	-	nd	-	-	-	+	-	nd	-	+
<i>P. sabulinigri</i> 13R26	E	-	+	-	+	-	-	-	-	-	-	-	-	-	+	-	nd	-	+
<i>P. sabulinigri</i> 13R41	E	+	++	-	-	-	-	-	-	-	nd	-	-	-	+	-	nd	-	-
<i>P. sabulinigri</i> PA4	B	+	+	+	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
<i>P. sabulinigri</i> PA20	B	+	++	+	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
<i>P. sabulinigri</i> RZ50	B	+	++	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. sabulinigri</i> RZ39	B	+	+	+	-	-	-	-	-	-	-	-	+	-	-	+	+	+	+
<i>P. sabulinigri</i> RZ33	B	+	++	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. sabulinigri</i> RZ38	B	+	+	+	-	-	-	-	-	-	+	-	-	-	-	+	+	+	+
<i>P. sabulinigri</i> RZ95	B	+	++	+	-	-	+	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. sabulinigri</i> RZ27	B	+	++	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. koreensis</i> RZ66	B	+	++	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+
<i>P. koreensis</i> RZ112	B	+	+	-	-	-	+++	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. koreensis</i> RZ4	B	-	++	+	-	-	+++	-	+	-	-	-	-	-	-	+	+	+	+
<i>P. koreensis</i> RZ72	B	+	+	-	-	-	+	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. koreensis</i> RZ3	B	+	++	-	-	-	+	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. koreensis</i> RZ86	B	+	+	-	-	-	+	-	-	-	-	-	-	+	-	+	+	+	+
<i>P. koreensis</i> RZ31	B	+	+	-	-	-	+	+	-	-	-	-	+	-	-	+	+	+	+
<i>P. koreensis</i> RZ98	B	+	+	-	-	-	+++	+	-	-	-	-	-	+	-	+	+	+	+
<i>P. koreensis</i> RZ96	B	+	+	-	-	-	+++	-	-	-	-	-	-	+	-	+	+	+	+

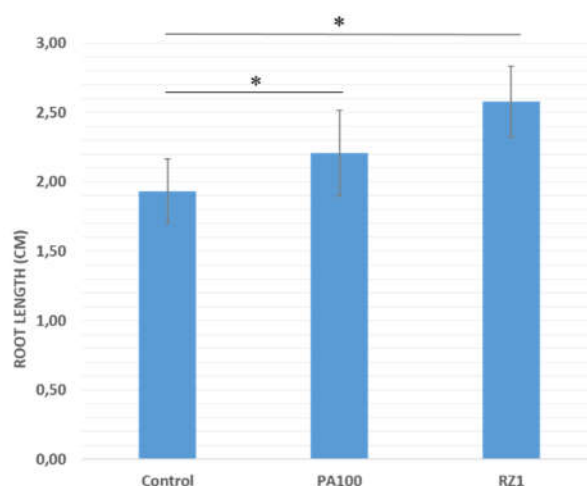
<i>P. koreensis</i> RZ81	B	-	+	-	-	-	+	-	-	-	-	-	-	+	-	+	+	+	+
<i>P. simiae</i> RZ105	B	+	++	-	-	-	+	-	-	-	-	-	-	+	-	+	+	+	+
<i>P. simiae</i> RZ85	B	+	+	+	-	-	+	-	-	-	-	-	-	+	-	+	+	+	+
<i>P. simiae</i> RZ109	B	+	++	-	-	-	++	-	+	-	-	-	+	+	-	+	+	+	+
<i>P. simiae</i> RZ110	B	-	+	-	-	-	++	-	-	-	-	-	-	+	-	+	+	+	+
<i>P. simiae</i> RZ1	B	+	+	+	-	-	-	-	-	+	+	+	-	+	-	+	+	+	+
<i>P. seleniipraecipitans</i> PA35	B	+	++	+	-	-	++	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. seleniipraecipitans</i> PA100	B	+	++	+	-	-	++	-	-	-	-	-	-	+	+	+	+	+	+
<i>P. guineae</i> RZ16	B	+	+	+	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+
<i>P. guineae</i> RZ7	B	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. fragi</i> RZ56	B	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. migulae</i> RZ17	B	+	++	+	-	-	++	-	-	-	-	-	-	-	-	+	+	+	+
<i>P. xanthomarina</i> 9PA12	B	+	++	+	-	-	-	-	-	-	-	-	-	-	nd	-	-	+	-
<i>Pseudomonas</i> sp. RZ32	B	+	+	-	-	-	++	-	-	-	-	-	+	-	-	+	+	+	+

P, phosphate solubilization; IAA, indole-3-acetic acid production; Sid, siderophores production; ACC, 1-aminocyclopropane-1-carboxylate deaminase production; *nifH*, nitrogen fixing gene presence; HCN, hydrogen cyanide production; *W.e.*, *Wautersia eutropha*; *A.t.*, *Agrobacterium tumefaciens*; *P.b.*, *Pantoea brenneri*; *P.e.*, *Pantoea eucalypti*; *E.r.*, *Erwinia rhapontici*; *C.f.*, *Curtobacterium flaccumfaciens*; Prot, protease; Lip, lipase; Cell, cellulase; Pect, pectinase; Amyl, amylase; Xyl, xylanase; +, tested positive; -, tested negative; ++, good producer; +++, strong producer; nd, not determined.

The production of cell wall lytic enzymes, such as proteases, pectinases, and lipases may be also involved in antimicrobial activity against pathogens, being beneficial to the plant and relevant in the endosphere (Compant et al., 2010; Pereira et al., 2015). Isolates were studied regarding their enzymatic activity. It was observed that most isolates produced xylanases (95 %), cellulases (87 %), pectinases (82 %), and amylases (82 %) (Table 3.4). Protease and lipase activities were observed in lower percentages of the isolates, 33 and 31 %, respectively. In addition to the importance to plant growth and health, endophytic *Pseudomonas* may also be interesting sources of enzymes for the industry.

### Effect of strain inoculation on *A. thaliana* root development

Due to their performance in terms of metal tolerance, PGP traits and antimicrobial and enzymatic activity, isolates PA100 (closely related to *P. seleniipraecipitans*), and RZ1 (affiliated with *P. simiae*) were selected to test their effect on *A. thaliana* root development. Both strains showed a positive effect on root development. Length of *A. thaliana* roots when inoculated with PA100 and RZ1, was on average 1.2 and 1.3 times higher than the length of roots of control plants (Figure 3.7). Differences between inoculated and control plants were significant for PA100 ( $p$ -value < 0.05) and RZ1 ( $p$ -value < 0.05). Strain RZ1 showed a significantly stronger positive effect on *A. thaliana* root development when compared to PA100 ( $p$ -value < 0.05).



**Figure 3.7** Effects of endophytic *Pseudomonas* inoculation on *A. thaliana* seedlings growth. Statistical significance is shown with  $p < 0.01$ (\*). Control, no inoculation; PA100 and RZ1 represent inoculation with isolates.

## Conclusions

This work highlights the diversity of endophytic *Pseudomonas* in *H. portulacoides* from metal(loid)-contaminated sites and their potential to be applied as facilitators of phytoremediation processes. Metal(loid) tolerance was common among analyzed isolates. Antibiotic resistance traits were mostly related with intrinsic characteristics of the species, and evidence of co-selection by metal(loid)s was not gathered.

Traits with potential benefits to plants, such as the production of PGP, production of enzymes, and activity against phytopathogens, were also observed in endophytic *Pseudomonas*. For example, isolate PA100 was tolerant to five metal(loid)s (Hg, As, Cu, Cr, and Ni), produced three PGP traits, presented antimicrobial activity against three representatives of phytopathogenic bacteria, and produced all tested enzymes. Isolate RZ1 was tolerant to two metals, produced three PGP traits, four enzymes and showed antimicrobial activity against three representatives of plant pathogens. Preliminary results show that these isolates promote *A. thaliana* root elongation. As demonstrated previously, bacterial strains may be used to inoculate plants in order to promote growth and/or help detoxify the environment (Rajkumar et al., 2009), thus potentially contributing for phytoremediation efficacy. The potential of the isolates characterized in this study for phytoremediation should be further tested *in vivo* in *H. portulacoides*.

## Acknowledgements

This work was supported by European Funds (FEDER) through COMPETE and by National Funds through the Portuguese Foundation for Science and Technology (FCT) within project PhytoMarsh (PTDC/AAC-651 AMB/118873/2010-FCOMP-01-0124-FEDER-019328). Authors also acknowledge FCT financing to CESAM (UID/AMB/50017/2013) and iBiMED (UID/BIM/04501/2013), Artur Alves (FCT Investigator Programme–IF/00835/2013), Isabel Henriques (FCT Investigator Programme–IF/00492/2013) and Cátia Fidalgo (PhD grant SFRH/BD/85423/2012). The authors wish to thank Sofia Pereira and Paula Castro (Universidade Católica Portuguesa, Portugal) for providing positive controls for plant growth promotion traits screening, Kornelia Smalla (Julius Kuhn Institut, Germany) for the phytopathogenic strains used in antimicrobial activity assays.

## References

- Aguilar-Barajas, E., Ramírez-Díaz, M. I., Riveros-Rosas, H. & Cervantes, C. (2010).** Heavy metal resistance in *Pseudomonads*. In: Ramos J, Filloux A (Eds.). *Pseudomonas*, Volume 6, Springer, pp. 255–282.
- Ali, S., Charles, T. C. & Glick, B. R. (2014).** Amelioration of high salinity stress damage by plant growth-promoting bacterial endophytes that contain ACC deaminase. *Plant Physiol Biochem* **80**, 160–167.
- Almeida, C. M. R., Dias, A. C., Mucha, A. P., Bordalo, A. A. & Vasconcelos, M. T. S. D. (2009).** Influence of surfactants on the Cu phytoremediation potential of a salt marsh plant. *Chemosphere* **75**, 135–140.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990).** Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Ando, S., Goto, M., Meunchang, S., Thongra-ar, P., Fujiwara, T., Hayashi, H. & Yoneyama, T. (2005).** Detection of *nifH* sequences in sugarcane (*Saccharum officinarum* L.) and pineapple (*Ananas comosus* [L.] Merr.). *Soil Sci Plant Nutr* **51**, 303–308.
- Anjum, N. A., Ahmad, I., Válega, M., Pacheco, M., Figueira, E., Duarte, A. C. & Pereira, E. (2011).** Impact of seasonal fluctuations on the sediment-mercury, its accumulation and partitioning in *Halimione portulacoides* and *Juncus maritimus* collected from Ria de Aveiro coastal lagoon (Portugal). *Water Air Soil Pollut* **222**, 1–15.
- Anyakora, C., Ehianeta, T. & Umukoro, O. (2013).** Heavy metal levels in soil samples from highly industrialized Lagos environment. *Afr J Environ Sci Technol* **7**, 917–924.
- Babu, A. G., Shea, P. J., Sudhakar, D., Jung, I. & Oh, B. (2015).** Potential use of *Pseudomonas koreensis* AGB-1 in association with *Miscanthus sinensis* to remediate heavy metal(loid)-contaminated mining site soil. *J Environ Manag* **151**, 160–166.
- Baker-Austin, C., Wright, M. S., Stepanauskas, R. & McArthur, J. V. (2006).** Co-selection of antibiotic and metal resistance. *Trends Microbiol* **14**, 176–182.
- Cambrollé, J., Mancilla-Leytón, J. M., Muñoz-Vallés, S., Luque, T. & Figueroa, M. T. (2012a).** Zinc tolerance and accumulation in the salt-marsh shrub *Halimione portulacoides*. *Chemosphere* **86**, 867–874.
- Cambrollé, J., Mancilla-Leytón, J. M., Muñoz-Vallés, S., Luque, T. & Figueroa, M. E. (2012b).** Tolerance and accumulation of copper in the salt-marsh shrub *Halimione portulacoides*. *Mar Pollut Bull* **64**, 721–728.
- Carattoli, A., Bertini, A., Villa, L., Falbo, V., Hopkins, K. L. & Threlfall, E. J. (2005).** Identification of plasmids by PCR-based replicon typing. *J Microbiol methods* **63**, 219–228.
- Carvalho, P. N., Basto, M. C. P., Silva, M. F. G. M., Machado, A., Bordalo, A. A. & Vasconcelos, M. T. S. D. (2010).** Ability of salt marsh plants for TBT remediation in sediments. *Environ Sci Pollut R* **17**, 1279–1286.



- Chaturvedi, S., Chandra, R. & Rai, V. (2006).** Isolation and characterization of *Phragmites australis* (L.) rhizosphere bacteria from contaminated site for bioremediation of colored distillery effluent. *Ecol Eng* **27**, 202–207.
- Chauhan, H., Bagyaraj, D. J., Selvakumar, G. & Sundaram, S. P. (2015).** Novel plant growth promoting rhizobacteria—prospects and potential. *Appl Soil Ecol* **95**, 38–53.
- CLSI (2012).** Performance standard for antimicrobial susceptibility testing—document approved standard M100-S22. CLSI, Wayne.
- Compant, S., Clément, C. & Sessitsch, A. (2010).** Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. *Soil Biol Biochem* **42**, 669–678.
- Costa, C. & Jesus-Rydin, C. (2001).** Site investigation on heavy metals contaminated ground in Estarreja - Portugal. *Eng Geol* **60**, 39–47.
- Couto, M. N. P. F. S., Basto, M. C. P. & Vasconcelos, M. T. S. D. (2011).** Suitability of different salt marsh plants for petroleum hydrocarbons remediation. *Chemosphere* **84**, 1052–1057.
- Deredjian, A., Colinon, C., Brothier, E., Favre-Bonté, S., Cournoyer, B. & Nazaret, S. (2011).** Antibiotic and metal resistance among hospital and outdoor strains of *Pseudomonas aeruginosa*. *Res Microbiol* **162**, 689–700.
- Duarte, B., Delgado, M. & Caçador, I. (2007).** The role of citric acid in cadmium and nickel uptake and translocation, in *Halimione portulacoides*. *Chemosphere* **69**, 836–840.
- Dworkin, M. & Foster, J. W. (1958).** Experiments with some microorganisms which utilize ethane and hydrogen. *J Bacteriol* **75**, 592–603.
- Fidalgo, C., Henriques, I., Rocha, J., Tacão, M. & Alves, A. (2016).** Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization and influence of metal(loid) contamination. *Environ Sci Pollut R* **23**, 10200–10214.
- Gaby, J. C. & Buckley, D. H. (2012).** A comprehensive evaluation of PCR primers to amplify the *nifH* gene of nitrogenase. *PLoS ONE* **7**, e42149.
- Ganeshan, G. & Kumar, A. M. (2005).** *Pseudomonas fluorescens*, a potential bacterial antagonist to control plant diseases. *J Plant Interact* **1**, 123–134.
- Gordon, S. A. & Weber, R. P. (1951).** Colorimetric estimation of indoleacetic acid. *Plant Physiol* **26**, 192–195.
- Kim, K., Roh, S. W., Chang, H., Nam, Y., Yoon, J., Jeon, C. O., Oh, H. & Bae, J. (2009).** *Pseudomonas sabulinigri* sp. nov., isolated from black beach sand. *Int J Syst Evol Microbiol* **59**, 38–41.

- Kim, O., Cho, Y., Lee, K., Yoon, S., Kim, M., Na, H., Park, S., Jeon, Y., Lee, J., Yi, H., Won, S. & Chun, J. (2012).** Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* **62**, 716–721.
- Livermore, D. M. (2001).** Of *Pseudomonas*, porins, pumps and carbapenems. *J Antimicrob Chemother* **47**, 247–250.
- Long, H. H., Schmidt, D. D. & Baldwin, I. T. (2008).** Native bacterial endophytes promote host growth in a species-specific manner; phytohormone manipulations do not result in common growth responses. *PLoS One* **3**, e2702.
- Lopez-Velasco, G., Tydings, H. A., Boyer, R. R., Falkinham, J. O. & Ponder, M. A. (2012).** Characterization of interactions between *Escherichia coli* O157:H7 with epiphytic bacteria in vitro and on spinach leaf surfaces. *Int J Food Microbiol* **153**, 351–357.
- Ma, Y., Prasad, M. N. V., Rajkumar, M. & Freitas, H. (2011).** Plant growth promoting rhizobacteria and endophytes accelerate phytoremediation of metalliferous soils. *Biotechnol Adv* **29**, 248–258.
- Malik, A. & Aleem, A. (2011).** Incidence of metal and antibiotic resistance in *Pseudomonas* spp. from the river water, agricultural soil irrigated with wastewater and ground water. *Environ Monit Assess* **178**, 293–308.
- Malik, A. & Jaiswal, R. (2000).** Metal resistance in *Pseudomonas* strains isolated from soil treated with industrial wastewater. *World J Microbiol Biot* **16**, 177–182.
- Martins, V. (2011).** Comunidade bacteriana endofítica cultivável de *Halimione portulacoides*. Dissertation, University of Aveiro. <http://ria.ua.pt/handle/10773/7377>
- Mesaros, N., Nordmann, P., Plésia, P., Roussel-Delvallez, M., Van Eldere, J., Glupczynski, Y., Van Laethem, Y., Jacobs, F., Lebecque, P., Malfroot, A., Tulkens, P. M. & Van Bambeke, F. (2007).** *Pseudomonas aeruginosa*: resistance and therapeutic options at the turn of the new millennium. *Clin Microbiol Infect* **13**, 560–578.
- Mucha, A. P., Almeida, C. M. R., Magalhães, C. M., Vasconcelos, M. T. S. D. & Bordalo, A. A. (2011).** Salt marsh plant-microorganism interaction in the presence of mixed contamination. *Int Biodeter Biodegr* **65**, 326–333.
- Nautiyal, C. S. (1999).** An efficient microbiological growth medium for screening phosphate solubilizing microorganisms. *FEMS Microbiol Lett* **170**, 265–270.
- Nies, D. H. (1999).** Microbial heavy-metal resistance. *Appl Microbiol Biot* **51**, 730–750.
- Penrose, D. M. & Glick, B. R. (2003).** Methods for isolating and characterizing ACC deaminase-containing plant growth-promoting rhizobacteria. *Physiol Plant* **118**, 10–15.

- Pereira, S. I. A., Barbosa, L. & Castro, P. M. L. (2015).** Rhizobacteria isolated from a metal-polluted area enhance plant growth in zinc and cadmium-contaminated soil. *Int J Environ Sci Technol* **12**, 2127–2142.
- Pereira, S. I. A., Monteiro, C., Vega, A. L. & Castro, P. M. L. (2016).** Endophytic culturable bacteria colonizing *Lavandula dentata* L. plants: isolation, characterization and evaluation of their plant growth-promoting activities. *Ecol Eng* **87**, 91–97.
- Pérez-Miranda, S. & Fernández, F. J. (2013).** Siderophores: what are they, and how are they detected? In: Amaya MG, Pacheco SV (eds) The struggle for iron: pathogen vs host. Cinvestav, Mexico.
- Pérez-Miranda, S., Cabirol, N., George-Téllez, R., Zamudio-Rivera, L. S. & Fernández, F. J. (2007).** O-CAS, a fast and universal method for siderophore detection. *J Microbiol methods* **70**, 127–131.
- Petatán-Sagahón, I., Anducho-Reyes, M. A., Silva-Rojas, H. V., Arana-Cuenca, A., Tellez-Jurado, A., Cárdenas-Álvarez, I. O. & Mercado-Flores, Y. (2011).** Isolation of bacteria with antifungal activity against the phytopathogenic fungi *Stenocarpella maydis* and *Stenocarpella macrospora*. *Int J Mol Sci* **12**, 5522–5537.
- Poole, K. (2011).** *Pseudomonas aeruginosa*: resistance to the max. *Front Microbiol* **2**, 1–13.
- Prado, S., Montes, J., Romalde, J. L. & Barja, J. L. (2009).** Inhibitory activity of *Phaeobacter* strains against aquaculture pathogenic bacteria. *Int Microbiol* **12**, 107–114.
- Preston, G. M. (2004).** Plant perceptions of plant growth-promoting *Pseudomonas*. *Philos T R Soc Lon B* **359**, 907–918.
- Proença, D. N., Francisco, R., Santos, C. V., Lopes, A., Fonseca, L., Abrantes, I. M. O. & Morais, P. V. (2010).** Diversity of bacteria associated with *Bursaphelenchus xylophilus* and other nematodes isolated from *Pinus pinaster* trees with pine wilt disease. *PLoS One* **5**, e15191.
- Rajkumar, M., Ae, N. & Freitas, H. (2009).** Endophytic bacteria and their potential to enhance heavy metal phytoextraction. *Chemosphere* **77**, 153–160.
- Reetha, A. K., Pavani, S. L. & Mohan, S. (2014).** Hydrogen cyanide production ability by bacterial antagonist and their antibiotics inhibition potential on *Macrophomina phaseolina* (Tassi.) Goid. *Int J Curr Microbiol Appl Sci* **3**, 172–178.
- R Core Team (2014).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Ryan, R. P., Germaine, K., Franks, A., Ryan, D. J. & Dowling, D. N. (2008).** Bacterial endophytes: recent developments and applications. *FEMS Microbiol Lett* **278**, 1–9.
- Shamim, S. & Rehman, A. (2013).** Antioxidative enzyme profiling and biosorption ability of *Cupriavidus metallidurans* CH34 and *Pseudomonas putida* mt2 under cadmium stress. *J Basic Microbiol* **55**, 374–381.

- Shi, C. L., Park, H. B., Lee, J. S., Ryu, S. & Ryu, C. M. (2010).** Inhibition of primary roots and stimulation of lateral root development in *Arabidopsis thaliana* by the rhizobacterium *Serratia marcescens* 90–166 is through both auxin-dependent and -independent signaling pathways. *Mol Cells* **29**, 251–258.
- Singh, P. & Cameotra, S. S. (2004).** Enhancement of metal bioremediation by use of microbial surfactants. *Biochem Biophys Res Commun* **319**, 291–297.
- Sousa, A. I., Caçador, I., Lillebø, A. I. & Pardal, M. A. (2008).** Heavy metal accumulation in *Halimione portulacoides*: intra- and extra-cellular metal binding sites. *Chemosphere* **70**, 850–857.
- Spiers, A. J., Buckling, A. & Rainey, P. B. (2000).** The causes of *Pseudomonas* diversity. *Microbiology* **146**, 2345–2350.
- Stout, L. & Nüsslein, K. (2010).** Biotechnological potential of aquatic plantmicrobe interactions. *Curr Opin Biotechnol* **21**, 339–345.
- Tacão, M., Moura, A., Correia, A. & Henriques, I. (2014).** Co-resistance to different classes of antibiotics among ESBL-producers from aquatic systems. *Water Res* **48**, 100–107.
- Tacão, M., Correia, A. & Henriques, I. S. (2015).** Low prevalence of carbapenem-resistant bacteria in river water: resistance is mostly related to intrinsic mechanisms. *Microb Drug Resist* **21**, 497–506.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013).** MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- Tarkowski, P. & Vereecke, D. (2014).** Threats and opportunities of plant pathogenic bacteria. *Biotechnol Adv* **32**, 215–229.
- Toribio, J., Escalante, A. E., Caballero-Mellado, J., González-González, A., Zavala, S., Souza, V. & Soberón-Chávez, G. (2011).** Characterization of a novel biosurfactant producing *Pseudomonas koreensis* lineage that is endemic to Cuatro Ciénegas Basin. *Syst Appl Microbiol* **34**, 531–535.
- Ullah, A., Heng, S., Munis, M. F. H., Fhad, S. & Yang, X. (2015).** Phytoremediation of heavy metals assisted by plant growth promoting (PGP) bacteria: a review. *Environ Exp Bot* **117**, 28–40.
- Válega, M., Lillebø, A. I., Pereira, M. E., Caçador, I., Duarte, A. C. & Pardal, M. A. (2008a).** Mercury in salt marshes ecosystems: *Halimione portulacoides* as biomonitor. *Chemosphere* **73**, 1224–1229.
- Válega, M., Lillebø, A. I., Pereira, M. E., Duarte, A. C. & Pardal, M. A. (2008b).** Longterm effects of mercury in a salt marsh: hysteresis in the distribution of vegetation following recovery from contamination. *Chemosphere* **71**, 765–772.
- Weyens, N., van der Lelie, D., Taghavi, S. & Vangronsveld, J. (2009).** Phytoremediation: plant-endophyte partnerships take the challenge. *Curr Opin Biotechnol* **20**, 248–254.

**Wu, X., Monchy, S., Taghavi, S., Zhu, W., Ramos, J. & van der Lelie, D. (2011).** Comparative genomics and functional analysis of niche-specific adaptation in *Pseudomonas putida*. *FEMS Microbiol Rev* **35**, 299–323.

**Wuana, R. A. & Okieimen, F. E. (2011).** Heavy metals in contaminated soils: a review of sources, chemistry, risks and best available strategies for remediation. *ISRN Ecol* **2011**, 1–20.

**Xu, Y., Zhou, Y., Ruan, J., Xu, S., Gu, J., Huang, S., Zheng, L., Yuan, B. & Wen, L. (2015).** Endogenous nitric oxide in *Pseudomonas fluorescens* ZY2 as mediator against the combined exposure to zinc and cefradine. *Ecotoxicology* **24**, 835–843.

**Yamamoto, S., Kasai, H., Arnold, D. L., Jackson, R. W., Vivian, A. & Harayama, S. (2000).** Phylogeny of the genus *Pseudomonas*: intrageneric structure reconstructed from the nucleotide sequences of *gyrB* and *rpoD* genes. *Microbiology* **146**, 2385–2394.

## Supplementary Material

**Supplementary Table S3.4** PCR primers and conditions for antibiotic and metal resistance genes amplification.

Target	Primer Sequence (5'- 3')	Amplicon size (bp)	Annealing temperature (°C)	Reference
<i>bla<sub>TEM</sub></i>	TEM_F: AAAGATGCTGAAGATCA TEM_R: TTTGGTATGGCTTCATTC	425	44	Speldooren et al., 1998
<i>bla<sub>SHV</sub></i>	SHV_F: GCGAAAGCCAGCTGTCGGGC SHV_R: GATTGGCGGCGCTGTTATCGC	304	62	Henriques et al., 2006a
<i>bla<sub>CTX-M</sub></i>	CTX_F: SCVATGTGCAGYACCAAGTAA CTX_R: GCTGCCGGTYTTATCVCC	652	55	Lu et al., 2010
<i>bla<sub>IMP</sub></i>	IMP_F: GAATAGAGTGGCTTAATTGTC IMP_R: GGTTTAAYAAAAACAACCACC	232	55	Henriques et al., 2006a
<i>bla<sub>VIM</sub></i>	VIM_F: GATGGTGTTTGGTCGCATATCG VIM_R: GCCACGTTCCCCGCAGACG	475	58	Henriques et al., 2006a
<i>bla<sub>KPC</sub></i>	KPC_F: CATTCAAGGGCTTTCTTGCTGC KPC_R: ACGACGGCATAGTCATTT	538	55	Dallenne et al., 2010
<i>bla<sub>GES</sub></i>	GES_F: AGTCGGCTAGACCGGAAAG GES_R: TTTGTCCGTGCTCAGGAT	399	57	Dallenne et al., 2010
<i>bla<sub>OXA-1-like</sub></i>	OXA1_F: ACACAATACATATCAACTTCGC OXA1_R: AGTGTGTTTAGAATGGTGATC	814	53	Ouellette et al., 1997
<i>bla<sub>OXA-2-like</sub></i>	OXA2_F: CAAGCCAAAGGCACGATAGTTG OXA2_R: CTAACCCATCCTACCCACC	561	56	Henriques et al., 2006a
<i>bla<sub>OXA-10-like</sub></i>	OXA10_F: CGTGCTTTGTAAGAGTAGCAG OXA10_R: CATGATTTTGGTGGGAATGG	652	53	Huovinen et al., 1988
<i>tet(A)</i>	tetA_F: GCTACATCCTGCCTTC tetA_R: GCATAGATCGGAAGAG	211	53	Nawaz et al., 2006
<i>tet(B)</i>	tetB_F: TCATTGCCGACCTCAG tetB_R: CCAACCATCACCATCC	391	53	Nawaz et al., 2006
<i>su1</i>	su1_F: CTGAACGATATCCAAGGATTYCC su1_R: AAAAATCCCACGGRTC	239	50	Heuer and Smalla, 2007
<i>Cat</i>	cat_F: CCTGCCACTCATCGCAGT cat_R: CCACCGTTGATATATCCC	623	60	Guerra et al., 2001
<i>int1</i>	int1_F: CCTCCCGCACGATGATC int1_R: TCCACGCATCGTCAGGC	280	55	Kraft et al., 1986
<i>merA</i>	merA_F: GTGCCGTCCAAGATCATGAT merA_R: TAGCCYACRGTSGCACATG	933	65	Deredjian et al., 2011

## References

- Cattoir, V., Poirel, L. & Nordmann, P. (2007).** Plasmid-mediated quinolone resistance determinant QnrB4 identified in France in an *Enterobacter cloacae* clinical isolate coexpressing a QnrS1 determinant. *Antimicrob Agents Ch* **51**, 2652-2653.
- Dallenne, C., Da Costa, A., Decré, D., Favier, C. & Arlet, G. (2010).** Development of a set of multiplex PCR assays for the detection of genes encoding important beta-lactamases in *Enterobacteriaceae*. *J Antimicrob Chemoth* **65**, 490-495.
- Guerra, B., Soto, S. M., Arguelles, J. M. & Mendoza, M. C. (2001).** Multidrug resistance is mediated by large plasmids carrying a class 1 integron in the emergent *Salmonella enterica* Serotype [4,5,12:i:2]. *Antimicrob Agents Ch* **45**, 1308-1308.
- Guillard, T., Moret, H., Brasmea, L., Carliera, A., Vernet-Garniera, V., Cambau, E. & de Champs, C. (2011).** Rapid detection of *qnr* and *qepA* plasmid-mediated quinolone resistance genes using real-time PCR. *Diagn Micr Infec Dis* **70**, 253-259.
- Henriques, I., Moura, A., Alves, A., Saavedra, M. J. & Correia, A. (2006a).** Analyzing diversity among  $\beta$ -lactamase encoding genes in aquatic environments. *FEMS Microbiol Ecol* **56**, 418–429.
- Henriques, I., Fonseca, F., Alves, A., Saavedra, M. J. & Correia, A. (2006b).** Occurrence and diversity of integrons and  $\beta$ -lactamase genes among ampicillin-resistant isolates from estuarine waters. *Res Microbiol* **157**, 938-947.
- Heuer, H. & Smalla, K. (2007).** Manure and sulfadiazine synergistically increased bacterial antibiotic resistance in soil over at least two months. *Environ Microbiol* **9**, 657-666.
- Huovinen, P., Huovinen, S. & Jacoby, G. A. (1988).** Sequence of PSE-2 $\beta$ -lactamase. *Antimicrob Agents Ch* **32**, 134–136.
- Kozak, G. K., Boerlin, P., Janecko, N., Reid-Smith, R. J. & Jardine, C. (2009).** Antimicrobial resistance in *Escherichia coli* isolates from swine and wild small mammals in the proximity of swine farms and in natural environments in Ontario, Canada. *Appl Environ Microbiol* **75**, 559-566.
- Lu, S., Zhang, Y., Geng, S., Li, T., Ye, Z., Zhang, D., Zou, F. & Zhou, H. (2010).** High diversity of extended spectrum beta-lactamase-producing bacteria in an urban river sediment habitat. *Appl Environ Microb* **76**, 5972-5976.
- Nawaz, M., Sung, K., Khan, S. A., Khan, A. A. & Steele, R. (2006).** Biochemical and molecular characterization of tetracycline-resistant *Aeromonas veronii* isolates from catfish. *Appl Environ Microb* **72**, 6461-6466.

**Ng, L. K., Martin, I., Alfa, M. & Mulvey, M. (2001).** Multiplex PCR for the detection of tetracycline resistant genes. *Mol Cell Probe* **15**, 209-215.

**Ouellette, M., Bissonnette, L. & Roy, P. H. (1997).** Precise insertion of antibiotic resistance determinants into Tn21-like transposons: Nucleotide sequence of the OXA-1 $\beta$ -lactamase gene. *P Natl Acad Sci USA* **84**, 7378–7382.

**Park, C. H., Robicsek, A., Jacoby, G. A., Sahm, D. & Hooper, D. C. (2006).** Prevalence in the United States of aac(6')-Ib-cr encoding a ciprofloxacin-modifying enzyme. *Antimicrob Agents Ch* **50**, 3953-3955.

**Poirel, L., Héritier, C., Tolün, V. & Nordmann, P. (2004).** Emergence of oxacillinase-mediated resistance to imipenem in *Klebsiella pneumoniae*. *Antimicrob Agents Ch* **48**, 15-22.

**Randall, L. P., Cooles, S. W., Osborn, M. K., Piddock, L. J. & Woodward, M. J. (2004).** Antibiotic resistance genes, integrons and multiple antibiotic resistance in thirty-five serotypes of *Salmonella enterica* isolated from humans and animals in the UK. *J Antimicrob Chemoth* **53**, 208–216.

**Speldooren, V., Heym, B., Labia, R. & Nicolas-Chanoine, M. H. (1998).** Discriminatory detection of inhibitor-resistant beta-lactamases in *Escherichia coli* by single-strand conformation polymorphism-PCR. *Antimicrob Agents Ch* **42**, 879-84.



## Chapter 4

---

Novel bacterial species

## Contents

**4.1** *Microbacterium diaminobutyricum* sp. nov., isolated from *Halimione portulacoides*, which contains diaminobutyric acid in its cell wall, and emended description of the genus *Microbacterium*

**4.2** *Saccharospirillum correiae* sp. nov., an endophytic bacterium isolated from the halophyte *Halimione portulacoides*

**4.3** *Altererythrobacter aveirensis* sp. nov. and *Altererythrobacter endophyticum* sp. nov., two endophytes from the salt marsh plant *Halimione portulacoides*

**4.4** *Zunongwangia endophytica* sp. nov., an endophyte isolated from the salt marsh plant *Halimione portulacoides*, and emended description of the genus *Zunongwangia*

**4.5** The endosphere of the salt marsh plant *Halimione portulacoides* is a diversity hotspot for the genus *Salinicola*: description of five novel species *Salinicola halimionae* sp. nov., *Salinicola aestuarina* sp. nov., *Salinicola endophytica* sp. nov., *Salinicola halophytica* sp. nov. and *Salinicola lusitana* sp. nov.

**4.1 *Microbacterium diaminobutyricum* sp. nov., isolated from *Halimione portulacoides*, which contains diaminobutyric acid in its cell wall, and emended description of the genus *Microbacterium***

**Authors**

Cátia Fidalgo,<sup>1,2</sup> Raúl Riesco,<sup>3</sup> Isabel Henriques,<sup>1,2</sup> Martha E. Trujillo<sup>3</sup> and Artur Alves<sup>1</sup>

<sup>1</sup> Departamento de Biologia, CESAM, Universidade de Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

<sup>2</sup> Departamento de Biologia, iBiMED, Universidade de Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

<sup>3</sup> Departamento de Microbiología y Genética, Edificio Departamental, Lab. 214, Campus Miguel de Unamuno, Universidad de Salamanca, 37007 Salamanca, Spain

**Publication status**

Published

Int J Syst Evol Microbiol

2016 Nov; 66(11):4492-4500

doi: 10.1099/ijsem.0.001379

## Abstract

Three actinobacterial strains were isolated from roots of the salt-marsh plant *Halimione portulacoides* collected in Ria de Aveiro, Portugal. Molecular typing using enterobacterial repetitive intergenic consensus ERIC-PCR fingerprinting showed the strains to be highly similar. Phylogenetic analyses based on the 16S rRNA gene sequence and multilocus sequence analysis (MLSA) using *gyrB*, *rpoB*, *recA* and *ppk* and 16S rRNA genes sequences showed that the strains represented a member of the genus *Microbacterium*, with *Microbacterium lacus* DSM 18910<sup>T</sup> as the closest phylogenetic relative. DNA–DNA hybridization between strain RZ63<sup>T</sup> and its closest relative was below 70 %, supporting the hypothesis that it represented a distinct genomic species. Chemotaxonomic analyses of the novel strains and their DNA G+C contents confirmed their affiliation to the genus *Microbacterium*, however, the peptidoglycan of RZ63<sup>T</sup> contained diaminobutyric acid as the diagnostic diamino acid. In addition, physiological and fatty acid analyses revealed differences between these strains and their phylogenetic relatives, reinforcing their status as a distinct species. Based on the physiological, genetic and chemotaxonomic characterization it is proposed that the strains studied represent a novel species of the genus *Microbacterium* for which the name *Microbacterium diaminobutyricum* sp. nov. is proposed (type strain RZ63<sup>T</sup>=DSM 27101<sup>T</sup>=CECT 8355<sup>T</sup>).

## Keywords

Actinobacteria, endophytic, salt-marsh, halophytes, taxonomy

## Main text

The genus *Microbacterium* includes a very large number of species (over 90) that have been isolated from a wide diversity of environments including plant-associated habitats. For example, *Microbacterium foliorum* and *Microbacterium phyllosphaerae* were isolated from the phyllosphere of grasses (Behrendt et al., 2001). Madhaiyan et al. (2010) described *Microbacterium azadirachtae*, a plant growth promoting bacterium, from the rhizoplane of *Azadirachta indica* seedlings. Karojet et al. (2012) described *Microbacterium yannicii* isolated from roots of *Arabidopsis thaliana*. In addition, endophytic strains of members of the genus *Microbacterium* have been isolated from different plants, such as sweet corn (*Zea mays*), cotton (*Gossypium hirsutum*) (McInroy & Kloepper, 1995), marigolds (*Tagetes erecta* and *T. patula*) (Sturz & Kimpinski, 2004), switchgrass (*Panicum virgatum*) (Gagne-Bourgue et al., 2013) and wild legumes (Zakhia et al., 2006).

The halophyte *Halimione portulacoides* (sea purslane) is a common and abundant species in salt marshes. This plant is known to sequester and tolerate high concentrations of toxic metals found in salt marsh sediments, namely mercury (Hg), and has been proposed as a bioindicator and biomonitor for Hg-contamination (Válega et al., 2008). Furthermore, this halophyte appears to be suitable for phytoremediation of Hg-contaminated salt marshes (Anjum et al., 2011).

It is recognised that endophytic bacteria establish complex, intimate and dynamic interactions with their plant hosts, having effects on fundamental aspects such as plant growth and health as well as promotion of stress resistance (Berg et al., 2014). A study of the culturable endophytic bacteria of *H. portulacoides* in the salt marsh Ria de Aveiro (Portugal) unveiled a phylogenetically and functionally diverse microbiota (Fidalgo et al., 2016). Among the 79 bacterial genera identified, *Microbacterium* arose as one of the most abundant and diverse especially in Hg-contaminated sites. Thus, the endosphere of *H. portulacoides* appears to represent a hotspot of diversity for the genus *Microbacterium*. This is corroborated by the recent description of three novel endophytic *Microbacterium* species (*M. endophyticum*, *M. halimionae* and *M. proteolyticum*) from *H. portulacoides* (Alves et al., 2014, 2015).

The aim of this study was to carry out a detailed taxonomic characterization of three novel isolates of endophytic *Microbacterium* isolates obtained from roots of the halophyte salt marsh plant *H. portulacoides* collected in the estuary Ria de Aveiro, Portugal, in the area of Laranjo Bay, Murtosa (40° 43' 48.0"N 8° 36' 45.5" W). Plant roots sterilization and bacterial isolation was performed as described previously (Fidalgo et al., 2016). Roots were washed with phosphate buffered saline (PBS; 140 mM NaCl, 2.5 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>) for 10 min, 96 % ethanol for 1 min, 5

% sodium hypochlorite for 30 min, 96 % ethanol for 1 min, and finally sterile distilled water for 3 min (repeated three times). In order to confirm sterility of root surface, aliquots (0.1 mL) of the final wash in sterile water were plated on tryptic soy agar (TSA) and incubated at 28 °C for 72 h. Plant roots were macerated in PBS and tenfold dilutions of the extracts were made on the same buffer. Aliquots of the dilutions (0.1 mL) were plated on TSA. After incubation at 28 °C for 72 h, the obtained colonies were transferred to new TSA plates for purification. Strains RZ63<sup>T</sup>, RZ102 and RZ104 were routinely cultured on TSA at 28 °C and preserved at -80 °C on tryptic soy broth (TSB; Merck) supplemented with 20 % (v/v) glycerol.

Genomic DNA was extracted using the Genomic DNA Purification Kit (Thermo Scientific) according to the manufacturer's instructions. Genetic relatedness of the isolates was evaluated by enterobacterial repetitive intergenic consensus ERIC-PCR fingerprinting using primers ERIC1 and ERIC2 (Alves et al., 2007). Cluster analysis of ERIC fingerprints was carried out with the software GelCompar II (Applied Maths).

The 16S rRNA gene was amplified using the universal primers 27F and 1492R (Lane, 1991). PCR products were purified with the DNA Clean & Concentrator 5 (Zymo Research) and sequenced at GATC Biotech (Cologne, Germany). The identification of phylogenetic neighbors was done by searching against the EzTaxon database (<http://www.ezbiocloud.net/eztaxon>; Kim et al., 2012). The 16S rRNA gene pairwise sequence similarities were determined using the program PHYDIT v3.1 (Chun, 2001).

A multilocus sequence analysis (MLSA) was performed using the scheme proposed by Richert et al. (2007). Thus, in addition to the 16S rRNA gene, the housekeeping genes coding for DNA gyrase subunit B (*gyrB*), RNA-polymerase subunit B (*rpoB*), recombinase A (*recA*) and polyphosphate kinase (*ppk*) were amplified and sequenced. A BLAST search of the primers used by Richert et al. (2007) against the GenBank database revealed a considerable number of mismatches in some primer binding sites of sequences from several species of the genus *Microbacterium*. Therefore, we developed novel sets of primers for the amplification and sequencing of *gyrB*, *rpoB*, *recA* and *ppk* genes in *Microbacterium* spp. (Table 4.1). For this, full-length sequences of these genes were retrieved from completed or draft genomes of *Microbacterium maritopicum* MF109 (accession number ATAO01000228), *Microbacterium barkeri* 2011-R4 (accession number AKVP01000049), *Microbacterium oleivorans* RIT293 (accession number JFY001000005), *Microbacterium paraoxydans* DH1b (accession number AYME01000002), *Microbacterium laevaniformans* OR221 (accession number AJGR01000511), *Microbacterium yannicii* PS01 (accession number

CAJF01000029), *Microbacterium testaceum* NS283 (accession number LDRU01000035), *Microbacterium luticocti* DSM 19459<sup>T</sup> (accession number AULS01000018), *Microbacterium indicum* DSM 19969<sup>T</sup> (accession number AULR01000004), and *Microbacterium gubbeenense* DSM 15944<sup>T</sup> (accession AUGQ01000021) available in GenBank. For each gene, sequences were aligned using CLUSTALX v2.1 (Larkin et al., 2007) and primers designed for regions showing high similarity. Primer features such as melting temperature (T<sub>m</sub>), DNA G+C content (%), self-annealing, GC clamp and hairpin formation were checked using the programs available at the Sequence Manipulation Suite site ([http://www.bioinformatics.org/sms2/pcr\\_primer\\_stats.html](http://www.bioinformatics.org/sms2/pcr_primer_stats.html)). PCR conditions for amplification of each gene are listed in Table 4.1.

**Table 4.1** Primer sets developed and PCR amplification conditions for each gene.

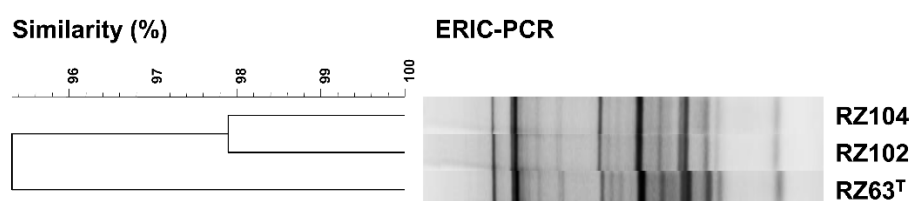
Gene	Primers sequence (5' – 3')	Amplicon size (approx.)	Amplification conditions	
			No. of cycles	Cycles
<i>gyrB</i>	gyrB F2 (AGA YSG CNT TCC TSA ACA AG)	750 bp	30 or 35	95 °C 1min
	gyrB R2 (CCT CSA CSA GGA ARA TCT CG) or			50-60 °C 30s
	gyrB R3 (CGC AGS GMS AGG ATC GCC TG)			72 °C 1min
<i>ppk</i>	ppk F1 (GCA ACC TCG ACG AGT TCT TC)	1650 bp	30	95 °C 1min
	ppk R1 (AGG TTG CGG TGC ATC ATG TC)			59-65 °C 30s
<i>recA</i>	recA F1 (CCN GAG TCN TCS GGT AAG AC)	700 bp	30	72 °C 1min or 2min
	or			95 °C 1min or 1min
	recA 6F (GGY CGC ATC RTC GAG ATC TAC)*			30s
	recA R1 (GMG TTC TCY TTG CCC TGR CC)			50-60 °C 30s
<i>rpoB</i>	rpoB F1 (GGC GAC TTC CCG CTS CAG AC)	1000 bp	30	72 °C 1min
	rpoB R1 (GGC ACA TSC GGC CGT AGT GC)			95 °C 1min
				60-65 °C 30s
				72 °C 1min or 1min
				15s

\*From Richert et al. (2007).

Nearly full-length 16S rRNA gene sequences and concatenated sequences of *gyrB*, *rpoB*, *recA* and *ppk* and 16S rRNA genes were aligned with sequences from type strains of *Microbacterium* species retrieved from the GenBank database using CLUSTALX v2.1 (Larkin et al., 2007). Phylogenetic analyses were performed with MEGA v6.0 (Tamura et al., 2013) using the neighbour-joining (NJ) (Saitou & Nei, 1987) and maximum-likelihood (ML) (Felsenstein, 1981) algorithms. Evolutionary distances were calculated using the Kimura-2-parameter model. *Clavibacter michiganensis* DSM

46364<sup>T</sup> was used as the outgroup taxon. A bootstrap analysis (Hillis & Bull, 1993) based on 1000 replicates was done to evaluate the robustness of the trees topologies.

ERIC-PCR fingerprinting showed that strains RZ63<sup>T</sup>, RZ102 and RZ104 had highly similar ERIC fingerprints (Figure 4.1). Nearly full-length 16S rRNA gene sequences (1436 bp) were obtained for all strains. 16S rRNA gene pairwise sequence comparisons showed that sequences from strains RZ63<sup>T</sup>, RZ102 and RZ104 shared 100 % similarity.



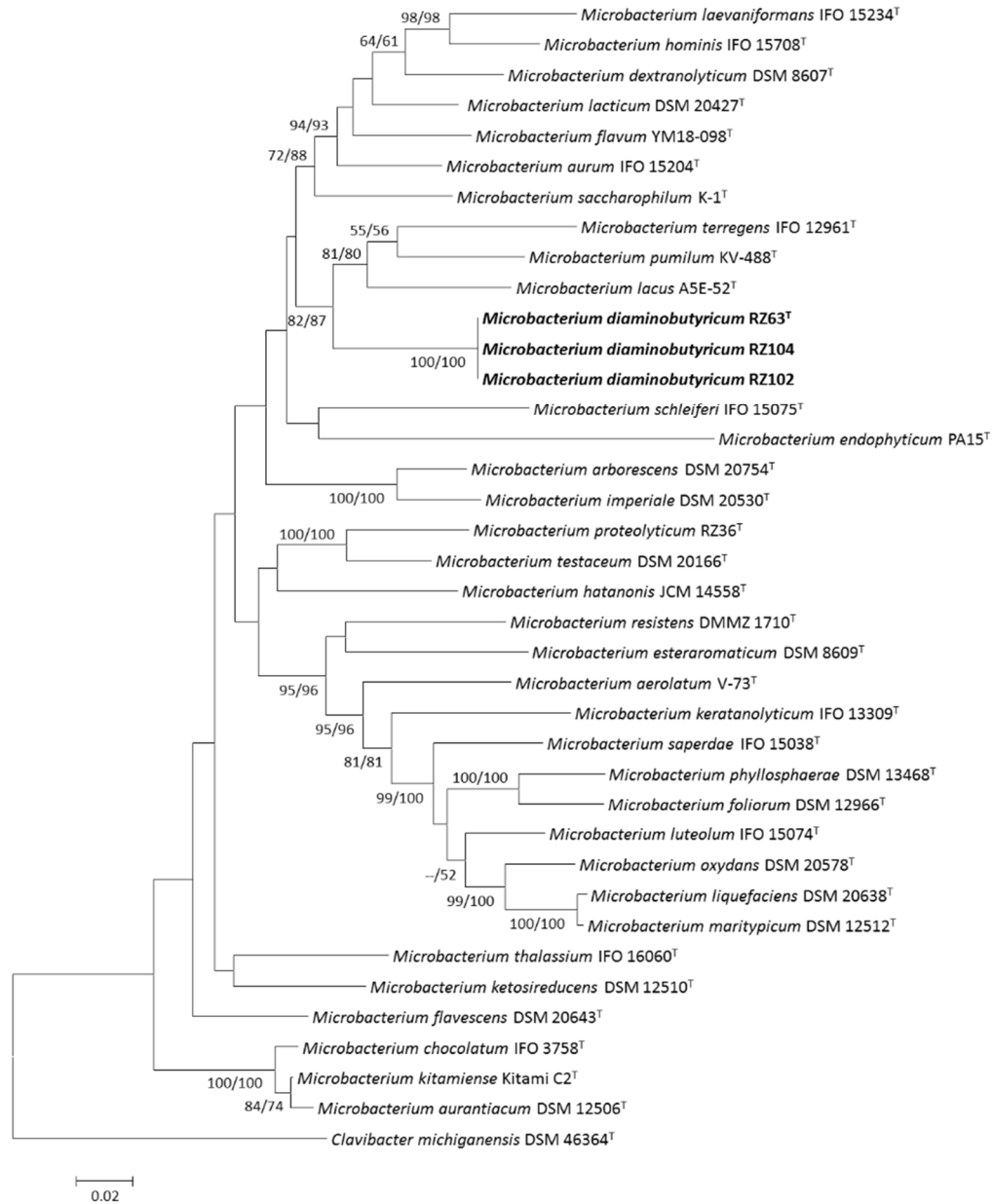
**Figure 4.1** UPGMA cluster analysis based on the Pearson coefficient of ERIC-PCR fingerprints of *Microbacterium* spp. strains isolated from roots of *Halimione portulacoides*. Similarity is indicated as a percentage.

A search against the EzTaxon database confirmed that all strains belonged to the genus *Microbacterium*. RZ63<sup>T</sup> and relatives showed the highest 16S rRNA gene sequence similarity to the type strain of *M. lacus* (98.9 %). Phylogenetic analyses of 16S rRNA gene sequences from all known species of the genus showed that these strains formed a distinct sub-clade (in both NJ and ML trees) with very high bootstrap support (99–100 %) within a larger clade containing several species of the genus *Microbacterium* (Supplementary Figure S4.1).

Using the primers developed, we amplified and sequenced the genes *gyrB*, *rpoB*, *recA* and *ppk* of RZ63<sup>T</sup>, RZ102 and RZ104, as well as of the type strains of *M. endophyticum* (PA15<sup>T</sup>), *M. hatanonis* (DSM 19179<sup>T</sup>), *M. lacus* (DSM 18910<sup>T</sup>), *M. proteolyticum* (RZ36<sup>T</sup>), *M. pumilum* (DSM 21018<sup>T</sup>) and *M. saccharophilum* (NCIMB 14782<sup>T</sup>) (Supplementary Table S4.1). For *M. halimionae* PA36<sup>T</sup>, all primer sets worked well except the ones developed for *gyrB*. Despite the design of several primer sets and the testing of different amplification conditions, it was impossible to amplify and sequence the *gyrB* gene. For this reason, *M. halimionae* PA36<sup>T</sup> was not included in the MLSA analyses. Phylogenetic analyses of concatenated sequences of *gyrB*, *rpoB*, *recA*, *ppk* and 16S rRNA genes confirmed the affiliation of the strains to the genus *Microbacterium*. RZ63<sup>T</sup> and its relatives formed a clearly



distinct clade with very high bootstrap support (100 %) (Figure 4.2). Moreover, this MLSA analysis confirmed that *M. lacus* DSM 18910<sup>T</sup> was the closest phylogenetic relative of the strains studied.



**Figure 4.2** Maximum-likelihood phylogenetic tree based on concatenated sequences of *gyrB*, *rpoB*, *recA* and *ppk* and 16S rRNA genes, showing the relationship between strains isolated from *Halimione portulacoides* and a set of type strains of species of the genus *Microbacterium*. *Clavibacter michiganensis* DSM 46364<sup>T</sup> was used as outgroup. ML/NJ bootstrap values ( $\geq 50$  %) from 1000 replicates are given at nodes. Bar, 0.02 substitutions per nucleotide position.

For morphological and physiological characterization, strains were cultured on TSA medium at 28 °C for 72 h. Growth at different temperatures (10, 15, 20, 25, 30, 37, 44 °C), NaCl concentrations (0.5, 1, 2, 4, 6, 8 %, w/v) and pH values (4.5, 5, 5.5, 6, 6.5, 8, 8.5, 9, 10, 11, 12) was tested using TSA as the basal medium. In the case of pH tests, the medium was adjusted using appropriate buffer solutions (McIlvaine, 1921; Bates & Bower, 1956). Oxidase and catalase activities were tested as described previously (Trujillo et al., 2006). Gram reaction was determined using a Gram Staining kit (Merck) according to the manufacturer's instructions. Physiological characteristics of the strains were determined using API 20NE, API ZYM and API 50CH test strips (Biomérieux) following the manufacturer's recommendations. Hydrolysis of casein and Tween 20 was evaluated by growing the strains on TSA supplemented with 1 % skim milk or Tween 20. After incubation at 28 °C for 72 h casein hydrolysis was revealed by the formation of a transparent halo around the colonies while hydrolysis of Tween 20 resulted in the formation of precipitates around the colonies. On the basis of pairwise comparisons of 16S rRNA gene sequences as well as MLSA analyses, *Microbacterium flavum* DSM 18909<sup>T</sup> and *Microbacterium lacus* DSM 18910<sup>T</sup> (Kageyama et al., 2007) were chosen as reference strains for comparative phenotypic studies.

The three endophytic strains produced circular, smooth and yellow-pigmented colonies on TSA medium. Cells were Gram-stain-positive, aerobic and short rods. RZ63<sup>T</sup> and related strains were oxidase-positive and catalase activity was variable (strain RZ102 was catalase-negative). None of the strains studied reduced nitrate to nitrite and all were negative for indole production. All strains (RZ63<sup>T</sup> and related strains) were able to grow in the presence of NaCl up to 8 % (w/v). Table 4.2 presents the differential characteristics between RZ63<sup>T</sup> (and related strains) and the type strains of *M. lacus* (DSM 18910<sup>T</sup>) and *M. flavum* (DSM 18909<sup>T</sup>).

Bacterial biomass for chemotaxonomic analyses was prepared by cultivating the strains in TSB at 28 °C for 72 h on a rotary shaker until reaching stationary phase. Cells were harvested by centrifugation and washed with 0.9 % (w/v) NaCl. The analyses of peptidoglycan structure were performed using TLC, 2D TLC, gas chromatography and GC/MS as described by Schumann (2011). Menaquinones were extracted from freeze-dried cells using the two-stage method described by Tindall (1990a, b), separated by TLC on silica gels and further analyzed by HPLC. For fatty acid methyl ester analysis, all strains were grown on TSA at 28 °C for 3 days or until good growth was obtained on the third quadrant (Sasser, 1990). Fatty acids were extracted, methylated and analyzed according to the standard protocol of the Sherlock Microbial Identification System (MIDI) and peaks were named using the database SACTIN6. Polar lipid and whole-cell hydrolysate sugar analyses

were carried out following the procedures of Minnikin et al. (1984) and Stanneck & Roberts (1974) using freeze-dried biomass.

**Table 4.2** Differential characteristics of strains obtained from *Halimione portulacoides* and phylogenetically related *Microbacterium* species, *M. flavum* DSM 18909<sup>T</sup> and *M. lacus* DSM 18910<sup>T</sup>.

Characteristic	RZ63 <sup>T</sup>	RZ102	RZ104	<i>M. flavum</i> DSM 18909 <sup>T</sup>	<i>M. lacus</i> DSM 18910 <sup>T</sup>
Growth at:					
Temperature range (°C)	10–37	10–37	10–37	10–37	15–30
pH range	6.5–11	6.5–11	6.5–11	6.5–11	6.0–11
NaCl range (w/v, %)	0.5–8	0.5–8	0.5–8	0.5–8	0.5–2
Oxidase	+	+	+	–	nd
Catalase	+	–	+	+	nd
D-glucose fermentation	–	+	+	+	–
Hydrolysis of:					
Gelatin	+	–	–	–	–
Casein	–	–	–	nd	nd
Aesculin	–	+	+	+	+
Tween 20	–	–	–	nd	nd
Urease	+	–	–	–	–
Esterase (C4)	–	+	+	+	+
Leucine arylamidase	–	+	+	+	+
Valine arylamidase	–	+	+	+	+
Cystine arylamidase	–	+	+	+	+
Trypsin	–	+	+	+	+
$\alpha$ -Galactosidase	–	+	w	–	+
$\beta$ -Galactosidase	+	+	+	+	+
$\beta$ -Glucuronidase	+	+	+	–	–
$\alpha$ -Mannosidase	–	+	+	–	w
Assimilation of:					
Glycerol	+	+	+	–	–
Erythritol	–	w	–	–	–
D-Arabinose	–	w	w	–	–
L-Arabinose	+	+	+	+	–
D-Ribose	–	w	w	–	–
D-Xylose	+	+	+	–	w
D-Galactose	+	+	+	+	–
L-Sorbose	–	w	w	–	–
L-Rhamnose	w	+	w	w	–
Methyl- $\alpha$ -D-glucopyranoside	+	–	w	–	–
N-acetylglucosamine	w	–	–	+	–
Amygdalin	+	+	+	+	–
Arbutin	+	+	+	+	–

Salicin	+	+	+	+	–
D-Lactose	+	+	+	+	–
D-Melibiose	–	–	–	+	–
D-Trehalose	+	+	+	+	w
Inulin	–	–	–	+	–
D-Melezitose	–	–	–	+	+
D-Raffinose	–	–	–	+	+
Starch	–	–	–	+	+
Glycogen	–	–	–	+	+
Xylitol	–	–	–	+	–
Gentiobiose	–	–	w	+	–
D-Lyxose	–	–	–	+	–
D-Tagatose	–	–	–	+	–
D-Fucose	–	–	–	+	–
L-Fucose	–	–	–	+	–
D-Arabitol	–	–	–	+	–
L-Arabitol	–	–	–	+	–
Potassium gluconate	w	–	w	+	–
Potassium 5-ketogluconate	–	–	–	w	–
Malic acid	–	–	–	+	–
Trisodium citrate	–	–	–	+	–
Phenylacetic acid	+	–	–	–	–
DNA G+C content (mol%)	68.9	nd	nd	70	69

+, positive; w, weak; –, negative; nd, not determined.

Chemotaxonomic analyses showed that peptidoglycan of strain RZ63<sup>T</sup> contained the amino acids alanine, glycine, homoserine (Hse), 2,4-diaminobutyric acid (DAB), glutamic acid and muramic acid (Mur) with molar ratios of 0.8 : 1.0 : 1.0 : 0.8 : 1.6 : 0.5. Partial hydrolysate of strain RZ63<sup>T</sup> contained the peptide Gly–Glu. These data suggest that the peptidoglycan of RZ63<sup>T</sup> belongs to type B2β with glycolyl residues (Schumann, 2011). In comparison to *M. lacus* DSM 18910<sup>T</sup> that has ornithine (Orn) as the diagnostic diamino acid, RZ63<sup>T</sup> has DAB as the diagnostic diamino acid. While several other members of the family *Microbacteriaceae* (e.g. members of the genera *Agromyces*, *Agrococcus*, *Clavibacter*, *Cryobacterium*, *Leucobacter* and *Plantibacter*) contain DAB as the diagnostic diamino acid (Evtushenko & Takeuchi, 2006), the strains of members of the genus *Microbacterium* described so far have either lysine (Lys) or Orn as the diagnostic diamino acid (Table 4.3). This is quite a distinguishing feature of RZ63<sup>T</sup>. Peptidoglycan analysis was repeated and the results were confirmed. Whole-cell sugar analyses revealed the presence of ribose, galactose, glucose and mannose.

Regarding menaquinones, RZ63<sup>T</sup> contained MK-10 (2 %), MK-11 (29 %), MK-12 (61 %) and MK-13 (6 %) and this composition is in accordance with that reported for members of the genus *Microbacterium*. The major fatty acids (>10 %) of RZ63<sup>T</sup> were anteiso-C<sub>15:0</sub> (69.1 %) and iso-C<sub>15:0</sub> (19.3 %) (Table 4.4). Quantitative differences of iso-C<sub>15:0</sub> and anteiso-C<sub>17:0</sub> were found between RZ63<sup>T</sup> and the remaining strains compared. Nevertheless, the tests were run again and the results were confirmed. With respect to the polar lipid profile of the novel strain (Supplementary Figure S4.2) the following lipids were detected: diphosphatidylglycerol, phosphatidylglycerol, one unidentified glycolipid, two unidentified phospholipids, four unidentified aminolipids and one unidentified lipid. These results are in agreement with the polar lipid profile reported for other species of the genus *Microbacterium*.

The DNA base composition of RZ63<sup>T</sup> was determined by HPLC (Mesbah et al., 1989). DNA was isolated by the modified procedure described by Gevers et al. (2001) and hybridizations were carried out at 53 °C in the presence of 50 % formamide using the modified microplate method (Ezaki et al., 1989; Goris et al., 1998; Cleenwerck et al., 2002). RZ63<sup>T</sup> was hybridized against the species with the highest 16S rRNA gene similarity, namely *M. lacus* DSM 18910<sup>T</sup>.

The DNA G+C content of RZ63<sup>T</sup> was 68.9 mol%. The value is within the range of 61–75 % reported for species of the genus *Microbacterium* (Suzuki & Hamada, 2012; Alves et al., 2014). DNA–DNA hybridization was used to clarify the relatedness of the strains and their closest relatives. DNA–DNA relatedness between RZ63<sup>T</sup> and *M. lacus* DSM 18910<sup>T</sup> was 14±1 %. Therefore, RZ63<sup>T</sup> represents a novel genomic species (Wayne et al., 1987).

On the basis of the phenotypic, chemotaxonomic and genotypic results described RZ63<sup>T</sup> (and related strains) are clearly distinct from their closest phylogenetic relative *M. lacus* DSM 18910<sup>T</sup> and represent a novel species in the genus *Microbacterium*, for which the name *Microbacterium diaminobutyricum* sp. nov. is proposed.

**Table 4.3** Chemotaxonomic markers of selected members of the genus *Microbacterium*.

	RZ63 <sup>T</sup>	RZ102	RZ104	<i>M. lacticum</i>	<i>M. flavum</i>	<i>M. lacus</i>	<i>M. halimionae</i>	<i>M. endophyticum</i>	<i>M. pumilum</i>	<i>M. terregens</i>	<i>M. proteolyticum</i>	<i>M. ginsengiterrae</i>
Main fatty acidsa	ai-C15 i-C15	ai-C15 i-C16 ai-C17	ai-C15 i-C16 ai-C17	ai-C15 ai-C17	ai-C15 ai-C17 i-C16	ai-C15 ai-C17 i-C16	ai-C15 ai-C17 i-C16	ai-C15 ai-C17 i-C16	ai-C15 i-C16 ai-C17	ai-C15 ai-C17	ai-C15 i-C16 ai-C17	ai-C15 ai-C17 i-C16
Cell-wall diamino acid	DAB	nd	nd	Lys	Lys	Orn	Orn	Orn	Orn	Orn	Orn	Orn
Polar lipid analysis	DPG PG GL 4AL 2PL 1L	nd	nd	DPG PG PGL PL	nd	nd	nd	nd	nd	DPG PG GL	DPG PG 2GL 3PL 1L	DPG PG GL 4AL
Whole cell-wall sugars	Gal Glu Man Rib	nd	nd	Gal Man Rha	nd	nd	nd	nd	Gal Rha	6dT Gal Rha	Gal Glu Rha Rib Xyl	Gal Rib Xyl
Menaquinones	MK-10 MK-11 MK-12 MK-13	nd	nd	MK-11 MK-12	MK-11 MK-12 MK-13	MK-11 MK-12 MK-13	MK-9 MK-10 MK-11 MK-12	MK-9 MK-10 MK-11 MK-12	MK-12 MK-13 MK-14	MK-12 MK-13	MK-10 MK-11 MK-12	MK-12 MK-13

ai-, anteiso-; i-, iso-; DAB, 2,4-diaminobutyric acid; Lys, lysine; Orn, ornithine; DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; GL, unidentified glycolipids; AL, unidentified amino lipid; PL, unidentified phospholipid; L, unidentified polar lipid; PGL, unidentified phosphoglycolipid; Gal, galactose; Glu, glucose; Man, mannose; Rib, ribose; Rha, rhamnose; 6dT, 6-deoxytalose; Xyl, xylose; nd, not determined.

**Table 4.4** Cellular fatty acid profiles of strains isolated from roots of *Halimione portulacoides* and phylogenetically related species of the genus *Microbacterium*.

Fatty acid	RZ63 <sup>T</sup>	RZ102	RZ104	<i>M. flavum</i> DSM 18909 <sup>T</sup>	<i>M. lacus</i> DSM 18910 <sup>T</sup>
anteiso-C <sub>13:0</sub>	1.8	0.2	0.2	–	0.1
iso-C <sub>14:0</sub>	3.8	1.0	0.9	0.7	0.8
C <sub>14:0</sub>	0.6	0.3	0.3	–	0.2
iso-C <sub>15:0</sub>	19.3	0.8	0.9	6.2	2.3
anteiso-C <sub>15:0</sub>	69.1	53.3	53.2	52.9	40.9
iso-C <sub>16:0</sub>	2.2	20.0	19.8	14.3	15.5
C <sub>16:0</sub>	0.4	3.9	4.2	1.5	4.1
anteiso-C <sub>17:0</sub>	0.9	19.7	19.9	20.7	33.4
iso-C <sub>17:0</sub>	–	0.1	0.1	2.7	1.5
C <sub>18:0</sub>	–	–	–	–	–
iso-C <sub>18:0</sub>	–	–	–	0.1	–

Values are percentages of total fatty acids; –, not detected.

### Emended description of the genus *Microbacterium* Orla-Jensen 1919 Takeuchi and Hatano 1998

The emended description of the genus is based on that given by Takeuchi & Hatano (1998) with the following modification: the peptidoglycan contains either lysine, ornithine or diaminobutyric acid as the diagnostic diamino acid.

#### Description of *Microbacterium diaminobutyricum* sp. nov.

*Microbacterium diaminobutyricum* (di.a.mi.no.bu.ty'ri.cum. N.L. n. acidum *diaminobutyricum*, containing 2,4-diaminobutyric acid in the peptidoglycan; N.L. neut. adj. *diaminobutyricum*, pertaining to diaminobutyric acid).

Cells are Gram-stain-positive, aerobic short rods. Colonies on TSA are circular, smooth and yellowish after 3 days incubation at 28 °C. The growth temperature range is 10–37 °C, with an optimum at 28 °C. Growth occurs at pH 6.5–11 with an optimum at pH 7.0, and in the presence of 0.5–8 % (w/v) NaCl with an optimum of 2 % (w/v). Cells are positive for oxidase activity. Catalase, urease, hydrolysis of aesculin and gelatin are variable. Nitrate reduction to nitrite, indole production and hydrolysis of casein and Tween 20 are not observed. The following substrates are used as sole carbon sources (API 50CH and API 20NE): glycerol, L-arabinose, D-xylose, D-galactose, D-glucose, D-

fructose, D-mannose, L-rhamnose, D-mannitol, amygdalin, arbutin, aesculin ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose, D-saccharose, D-trehalose and D-turanose; the utilization of erythritol, D-arabinose, D-ribose, L-sorbose, methyl- $\alpha$ -D-glucopyranoside, *N*-acetylglucosamine, gentiobiose, potassium gluconate and phenylacetic acid is variable; L-xylose, D-adonitol, methyl- $\beta$ -D-xylopyranoside, dulcitol, inositol, D-sorbitol, methyl- $\alpha$ -D-mannopyranoside, D-melibiose, inulin, D-melezitose, D-raffinose, starch, glycogen, xylitol, D-tagatose, D-lyxose. D-fucose, L-fucose, D-arabitol, L-arabitol, potassium 2-ketogluconate, potassium 5-ketogluconate, capric acid, adipic acid, malic acid and trisodium citrate are not used. In the API ZYM tests, cells are positive for esterase lipase (C8), lipase (C14),  $\alpha$ -chymotrypsin, alkaline phosphatase, acid phosphatase, naphthol-AS-BI-phosphohydrolase,  $\beta$ -galactosidase,  $\beta$ -glucuronidase,  $\alpha$ -glucosidase,  $\beta$ -glucosidase and *N*-acetyl-bglucosaminidase; results are variable for esterase (C4), leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin,  $\alpha$ -galactosidase and  $\alpha$ -mannosidase; and cells are negative for arginine dihydrolase and  $\alpha$ -fucosidase. The cell-wall peptidoglycan type is B2 $\beta$  with glycolyl residues and contains 2,4-diaminobutyric acid as diamino acid. The cell-wall sugars are ribose, mannose, glucose and galactose. The major menaquinones are MK-11 and MK-12. The major cellular fatty acids are anteiso-C<sub>15:0</sub> and iso-C<sub>15:0</sub> or anteiso-C<sub>15:0</sub>, anteiso-C<sub>17:0</sub> and iso-C<sub>16:0</sub>. The lipids comprise diphosphatidylglycerol, phosphatidylglycerol, one unidentified glycolipid, two unidentified phospholipids, four unidentified aminolipids and one unidentified lipid.

The type strain is RZ63<sup>T</sup> (=DSM 27101<sup>T</sup> =CECT 8355<sup>T</sup>), isolated from roots of *Halimione portulacoides* (sea purslane) collected in Ria de Aveiro, Laranjo Bay, Murtosa, Portugal (40° 43' 48.0" N 8° 36' 45.5" W). The G+C content of the genomic DNA of the type strain is 68.9 mol%.

## Acknowledgements

This work was supported by European Funds (FEDER) through COMPETE and by National Funds through the Portuguese Foundation for Science and Technology (FCT) within project PhytoMarsh (PTDC/AAC-AMB/118873/2010 – FCOMP-01-0124-FEDER-019328). FCT is acknowledged for financing to the Centre for Environmental and Marine Studies (CESAM) (UID/AMB/50017/2013), the institute for research in Biomedicine (iBiMED) (UID/BIM/04501/2013), Artur Alves (FCT Investigator Programme – IF/00835/2013), Isabel Henriques (FCT Investigator Programme – IF/00492/2013) and Cátia Fidalgo (PhD grant - SFRH/BD/85423/2012). M. E. Trujillo was funded by the Ministerio de Economía y Competitividad under project CGL2014-52735-P.



## References

- Alves, A., Phillips, A. J. L., Henriques, I. & Correia, A. (2007). Rapid differentiation of species of *Botryosphaeriaceae* by PCR fingerprinting. *Res Microbiol* **158**, 112–121.
- Alves, A., Correia, A., Igual, J. M. & Trujillo, M. E. (2014). *Microbacterium endophyticum* sp. nov. and *Microbacterium halimionae* sp. nov., endophytes isolated from the salt-marsh plant *Halimione portulacoides* and emended description of the genus *Microbacterium*. *Syst Appl Microbiol* **37**, 474–479.
- Alves, A., Riesco, R., Correia, A. & Trujillo, M. E. (2015). *Microbacterium proteolyticum* sp. nov. isolated from roots of *Halimione portulacoides*. *Int J Syst Evol Microbiol* **65**, 1794–1798.
- Anjum, N. A., Ahmad, I., Válega, M., Pacheco, M., Figueira, E., Duarte, A. C. & Pereira, E. (2011). Impact of seasonal fluctuations on the sediment-mercury, its accumulation and partitioning in *Halimione portulacoides* and *Juncus maritimus* collected from Ria de Aveiro coastal lagoon (Portugal). *Water, Air, Soil Pollut* **222**, 1–15.
- Bates, R. G. & Bower, V. E. (1956). Alkaline solutions for pH control. *Anal Chem* **28**(8), 1322–1324.
- Behrendt, U., Ulrich, A. & Schumann, P. (2001). Description of *Microbacterium foliorum* sp. nov. and *Microbacterium phyllosphaerae* sp. nov., isolated from the phyllosphere of grasses and the surface litter after mulching the sward, and reclassification of *Aureobacterium resistens* (Funke et al. 1998) as *Microbacterium resistens* comb. nov. *Int J Syst Evol Microbiol* **51**, 1267–1276.
- Berg, G., Grube, M., Schlöter, M., Smalla, K. (2014). Unraveling the plant microbiome: looking back and future perspectives. *Front Microbiol* **5**, 148.
- Cleenwerck, I., Vandemeulebroecke, K., Janssens, D. & Swings, J. (2002). Re-examination of the genus *Acetobacter*, with descriptions of *Acetobacter cerevisiae* sp. nov. and *Acetobacter malorum* sp. nov. *Int J Syst Evol Microbiol* **52**, 1551–1558.
- Chun, J. (2001). PHYDIT. Molecular sequence editor for phylogeny. Version 3.1. <http://plaza.snu.ac.kr/~jchun/phydit/>
- Evtushenko, L. I. & Takeuchi, M. (2006). The family *Microbacteriaceae*. In *The Prokaryotes: a Handbook on the Biology of Bacteria*, 3rd edn, vol. 3, pp. 1020–1098. Edited by M. Dworkin, S. Falkow, E. Roenigk, K. H. Schleifer & E. Stackebrandt. New York: Springer.
- Ezaki, T., Hashimoto, Y. & Yabuuchi, E. (1989). Fluorometric deoxyribonucleic acid-deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane filter hybridization in which radioisotopes are used to determine genetic relatedness among bacterial strains. *Int J Sys Bacteriol* **39**, 224–229.

- Felsenstein, J. (1981).** Evolutionary trees from DNA sequences: a maximum-likelihood approach. *J Mol Evol* **17**, 368–376.
- Fidalgo, C., Henriques, I., Rocha, J., Tacão, M. & Alves, A. (2016).** Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization and influence of metal(loid) contamination. *Environ Sci Pollut Res* **23**:10200–10214.
- Gagne-Bourgue, F., Aliferis, K. A., Seguin, P., Rani, M., Samson, R., & Jabaji, S. (2012).** Isolation and characterization of indigenous endophytic bacteria associated with leaves of switchgrass (*Panicum virgatum* L.) cultivars. *J Appl Microbiol* **114**, 836–853.
- Gevers, D., Huys, G. & Swings, J. (2001).** Application of rep-PCR fingerprinting for identification of *Lactobacillus* species. *FEMS Microbiol Lett* **205**, 31–36.
- Goris, J., Suzuki, K., De Vos, P., Nakase, T. & Kersters, K. (1998).** Evaluation of microplate DNA-DNA hybridization method compared with the initial renaturation method. *Can J Microbiol* **44**, 1148–1153.
- Hillis, D. M. & Bull, J. J. (1993).** An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst Biol* **42**, 182–192.
- Kageyama, A., Takahashi, Y., Matsuo, Y., Adachi, K., Kasai, H., Shizuri, Y., & Omura, S. (2007).** *Microbacterium flavum* sp. nov. and *Microbacterium lacus* sp. nov., isolated from marine environments. *Actinomycetologica* **21**, 53–58.
- Karojet, S., Kunz, S. & van Dongen, J. T. (2012).** *Microbacterium yannicii* sp. nov., isolated from *Arabidopsis thaliana* roots. *Int J Syst Evol Microbiol* **62**, 822–826.
- Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H., Yi, H., Won, S. & Chun, J. (2012).** Introducing EzTaxon: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* **62**, 716–721.
- Lane, D. J. (1991).** 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics*, pp. 115–175. Edited by E. Stackebrandt & M. Goodfellow. Chichester: Wiley.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007).** Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.
- Madhaiyan, M., Poonguzhali, S., Lee, J.-S., Lee, K.-C., Saravanan, V. S. & Santhanakrishnan, P. (2010).** *Microbacterium azadirachtae* sp. nov., a plant growth-promoting actinobacterium isolated from the rhizoplane of neem seedlings. *Int J Syst Evol Microbiol* **60**, 1687–1692.
- McIlvaine, T. C. (1921).** A buffer solution for colorimetric comparison. *J Biol Chem* **49**, 183–186.

- Mesbah, M., Premachandran, U. & Whitman, W. B. (1989).** Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* **39**, 159–167.
- McInroy, J. A. & Kloepper, J. W. (1995).** Survey of indigenous bacterial endophytes from cotton and sweet corn. *Plant Soil* **173**, 337–342.
- Minnikin, D. E., O'Donnell, A. G., Goodfellow, M., Alderson, G., Athalye, M., Schaal, K. & Parlett, J. H. (1984).** An integrated procedure for extracting bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods* **2**, 233–241.
- Richert, K., Brambilla, E. & Stackebrandt, E. (2007).** The phylogenetic significance of peptidoglycan types: molecular analysis of the genera *Microbacterium* and *Aureobacterium* based upon sequence comparison of *gyrB*, *rpoB*, *recA* and *ppk* and 16S rRNA genes. *Syst and Appl Microbiol* **30**, 102–108.
- Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Sasser, M. (1990).** Identification of bacteria by gas chromatography of cellular fatty acids. *USFCC News* **20**, 1–6.
- Schumann, P. (2011).** Peptidoglycan structure. *Methods Microbiol* **38**, 101–129.
- Staneck, J. L. & Roberts, G. D. (1974).** Simplified approach to the identification of aerobic actinomycetes by thin-layer chromatography. *Appl Microbiol* **28**, 226–231.
- Sturz, A. V. & Kimpinski, J. (2004).** Endoroot bacteria derived from marigolds (*Tagetes* spp.) can decrease soil population densities of root-lesion nematodes in the potato root zone. *Plant Soil* **262**, 241–249.
- Suzuki, K.-I., Hamada, M. (2012).** Genus I. *Microbacterium*, in: Goodfellow, M., Kämpfer, P., Busse, H.-J., Trujillo, M. E., Suzuki, K.-I., Ludwig, W., Whitman, W. B. (Eds.), *Bergey's Manual of Systematic Bacteriology*, vol. 5, second ed., Springer, New York, NY, pp. 814–852.
- Takeuchi, M. & Hatano, K. (1998).** Union of the genera *Microbacterium* Orla-Jensen and *Aureobacterium* Collins et al. in a redefined genus *Microbacterium*. *Int J Syst Bacteriol* **48**, 739–747.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013).** MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- Tindall, B. J. (1990a).** A comparative study of the lipid composition of *Halobacterium saccharovorum* from various sources. *Syst Appl Microbiol* **13**, 128–130.
- Tindall, B. J. (1990b).** Lipid composition of *Halobacterium lacusprofundi*. *FEMS Microbiol Lett* **66**, 199–202.

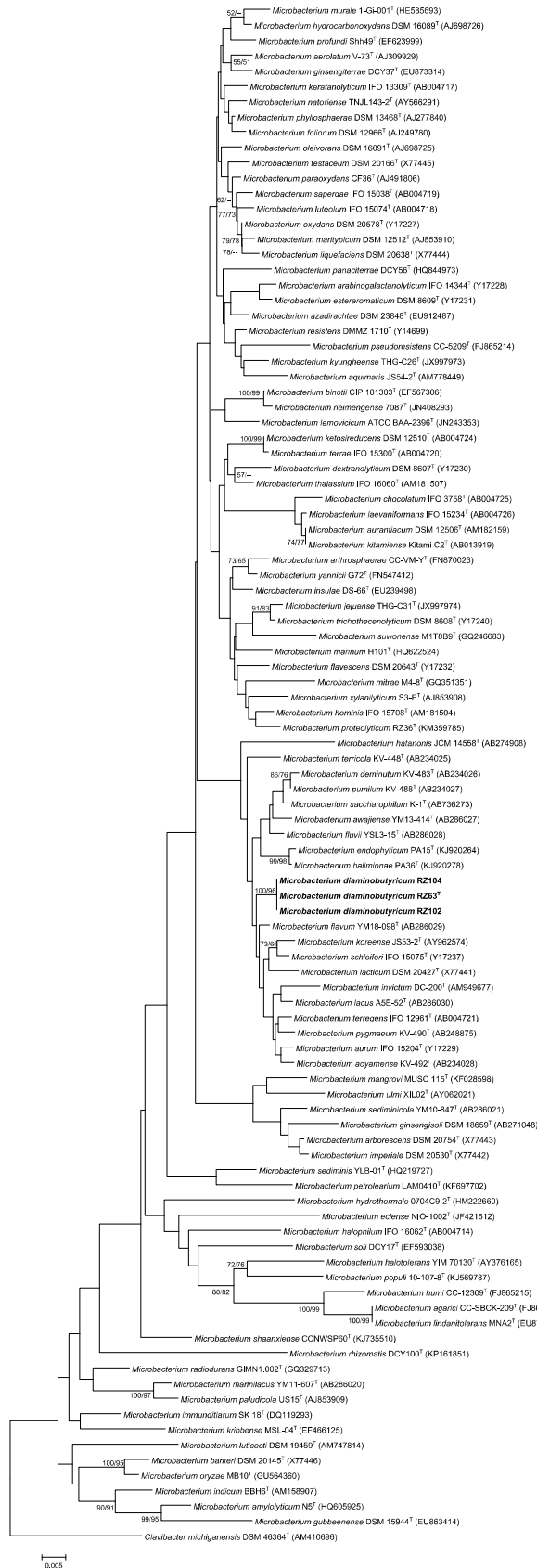
**Trujillo, M. E., Kroppenstedt, R. M., Schumann, P., Carro, L., & Martinez-Molina, E. (2006).** *Micromonospora coriariae* sp. nov., isolated from root nodules of *Coriaria myrtifolia*. *Int J Syst Evol Microbiol* **56**, 2381-2385.

**Válega, M., Lillebø, A. I., Pereira, M. E., Caçador, I., Duarte, A. C. & Pardal, M. A. (2008).** Mercury in salt marshes ecosystems: *Halimione portulacoides* as biomonitor. *Chemosphere* **73**, 1224-1229.

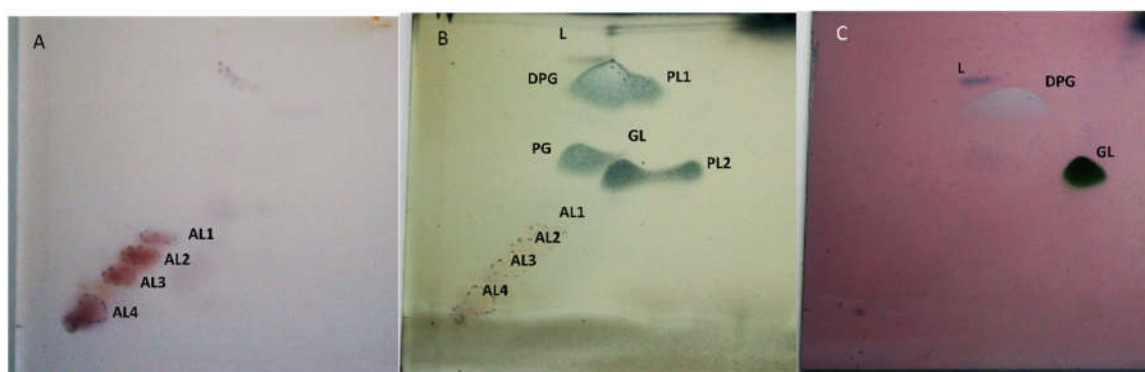
**Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., *et al.* (1987).** International Committee on Systematic Bacteriology Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* **37**, 463–464.

**Zakhia, F., Jeder, H., Willems, A., Gillis, M., Dreyfus B. & de Lajudie, P. (2006).** Diverse bacteria associated with root nodules of spontaneous legumes in Tunisia and first report for *nifH*-like gene within the genera *Microbacterium* and *Starkeya*. *Microb Ecol* **51**, 375–393.

## Supplementary material



**Supplementary Figure S4.1** Maximum-likelihood phylogenetic tree based on nearly full-length 16S rRNA gene sequences, showing the relationship between strains isolated from *Halimione portulacoides* and type strains of the genus *Microbacterium*. *Clavibacter michiganensis* DSM 46364<sup>T</sup> was used as outgroup. ML/NJ bootstrap values ( $\geq 50\%$ ) from 1000 replicates are given at nodes. Bar, 0.005 substitutions per nucleotide position.



**Supplementary Figure S4.2** Polar lipid profile of strain RZ63<sup>T</sup> after staining with: **A**, ninhydrin, **B**, molybdatophosphoric acid and **C**, anisaldehyde. DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; PL1 and PL2, unidentified phospholipids; GL, unknown glycolipid; AL1, AL2, AL3 and AL4, unidentified aminolipids; L, unidentified lipid.

**Supplementary Table S4.1** GenBank accession numbers of *gyrB*, *rpoB*, *recA*, and *ppk* genes.

Species	GenBank accession			
	<i>gyrB</i>	<i>ppk</i>	<i>recA</i>	<i>rpoB</i>
<i>Microbacterium aerolatum</i> V-73 <sup>T</sup>	AM181475	AM181537	AM181508	AM181564
<i>Microbacterium arborescens</i> DSM 20754 <sup>T</sup>	AM181476	AM181538	AM181509	AM181565
<i>Microbacterium aurantiacum</i> DSM 12506 <sup>T</sup>	AM181477	AM181539	AM181510	AM181566
<i>Microbacterium aurum</i> IFO 15204 <sup>T</sup>	AM181478	AM181540	AM181512	AM181567
<i>Microbacterium chocolatum</i> IFO 3758 <sup>T</sup>	AM181479	AM181541	AM181511	AM181568
<i>Microbacterium dextranolyticum</i> DSM 8607 <sup>T</sup>	AM181480	AM181542	AM181513	AM181569
<i>Microbacterium endophyticum</i> PA15 <sup>T</sup>	KU843508	KU843518	KU843528	KU843538
<i>Microbacterium esteraromaticum</i> DSM 8609 <sup>T</sup>	AM181481	AM181543	AM181514	AM181570
<i>Microbacterium flavescens</i> DSM 20643 <sup>T</sup>	AM181482	AM181544	AM181515	AM181571
<i>Microbacterium flavum</i> YM18-098 <sup>T</sup>	KU843509	KU843519	KU843529	KU843539
<i>Microbacterium foliorum</i> DSM 12966 <sup>T</sup>	AM181483	AM181545	AM181516	AM181572
<i>Microbacterium hatanonis</i> JCM 14558 <sup>T</sup>	KU843510	KU843520	KU843530	KU843540
<i>Microbacterium hominis</i> IFO 15708 <sup>T</sup>	AM181484	AM181546	AM181517	AM181573
<i>Microbacterium imperiale</i> DSM 20530 <sup>T</sup>	AJ784798	AM181547	AM181518	AM181574
<i>Microbacterium keratanolyticum</i> IFO 13309 <sup>T</sup>	AM181485	AM181548	AM181519	AM181575
<i>Microbacterium ketosireducens</i> DSM 12510 <sup>T</sup>	AM181486	AM181549	AM181520	AM181576
<i>Microbacterium kitamiense</i> Kitami C2 <sup>T</sup>	AM181487	AM181550	AM181521	AM181577
<i>Microbacterium lacticum</i> DSM 20427 <sup>T</sup>	AM181488	AM181551	AM181522	AM181578
<i>Microbacterium lacus</i> ASE-52 <sup>T</sup>	KU843511	KU843521	KU843531	KU843541
<i>Microbacterium laevaniformans</i> IFO 15234 <sup>T</sup>	AM181490	AM181553	AM181524	AM181580
<i>Microbacterium liquefaciens</i> DSM 20638 <sup>T</sup>	AM181489	AM181552	AM181523	AM181579

<i>Microbacterium luteolum</i> IFO 15074 <sup>T</sup>	AM181491	AM181554	AM181525	AM181581
<i>Microbacterium maritypicum</i> DSM 12512 <sup>T</sup>	AM181492	AM181555	AM181526	AM181582
<i>Microbacterium oxydans</i> DSM 20578 <sup>T</sup>	AM181493	AM181556	AM181527	AM181583
<i>Microbacterium phyllosphaerae</i> DSM 13468 <sup>T</sup>	AM181494	AM181557	AM181528	AM181584
<i>Microbacterium proteolyticum</i> RZ36 <sup>T</sup>	KU843512	KU843522	KU843532	KU843542
<i>Microbacterium pumilum</i> KV-488 <sup>T</sup>	KU843513	KU843523	KU843533	KU843543
<i>Microbacterium resistens</i> DMMZ 1710 <sup>T</sup>	AM181495	AM181558	AM181529	AM181585
<i>Microbacterium saccharophilum</i> K-1 <sup>T</sup>	KU843514	KU843524	KU843534	KU843544
<i>Microbacterium saperdae</i> IFO 15038 <sup>T</sup>	AM181496	AM181559	AM181530	AM181586
<i>Microbacterium schleiferi</i> IFO 15075 <sup>T</sup>	AM181497	AM181560	AM181531	AM181587
<i>Microbacterium terregens</i> IFO 12961 <sup>T</sup>	AM181498	AM181561	AM181532	AM181588
<i>Microbacterium testaceum</i> DSM 20166 <sup>T</sup>	AM181499	AM181562	AM181533	AM181589
<i>Microbacterium thalassium</i> IFO 16060 <sup>T</sup>	AM181500	AM181502	AM181534	AM181590
<i>Microbacterium</i> sp. RZ63 <sup>T</sup>	KU843515	KU843525	KU843535	KU843545
<i>Microbacterium</i> sp. RZ102	KU843516	KU843526	KU843536	KU843546
<i>Microbacterium</i> sp. RZ104	KU843517	KU843527	KU843537	KU843547
<i>Clavibacter michiganensis</i> DSM 46364 <sup>T</sup>	AM410850	AM410798	AM410824	AM410876

Sequences in italics were obtained in this study.

## **4.2 *Saccharospirillum correiae* sp. nov., an endophytic bacterium isolated from the halophyte *Halimione portulacoides***

### **Authors**

Cátia Fidalgo<sup>1</sup>, Jaqueline Rocha<sup>1</sup>, Diogo Neves Proença<sup>2</sup>, Paula V. Morais<sup>2,3</sup>, Artur Alves<sup>1</sup>, Isabel Henriques<sup>1</sup>

<sup>1</sup> Departamento de Biologia, CESAM, Universidade de Aveiro, Aveiro, Portugal

<sup>2</sup> CEMUC, University of Coimbra, 3030-788 Coimbra, Portugal

<sup>3</sup> Department of Life Sciences, FCTUC, University of Coimbra, 3000-456 Coimbra, Portugal

### **Publication status**

In press

Int J Syst Evol Microbiol

doi: 10.1099/ijsem.0.001914



## Abstract

A Gram-stain negative, oxidase- and catalase-positive, motile, aerobic, non-pigmented spirillum, designated CPA1<sup>T</sup>, was isolated from the surface-sterilized tissues of a halophyte, *Halimione portulacoides*, collected from a salt marsh in Aveiro, Portugal. The isolate was mesophilic, facultatively alkaliphilic and halophilic, and grew between 18 and 42.5 °C (optimum 30 °C), from pH 5.0 to 11.5 (optimum 7.0-7.5), from 0.5 to 5 % NaCl (w/v, optimum 2 %). Analysis of the 16S rRNA gene sequence showed that this strain belongs to the genus *Saccharospirillum*, as the highest sequence similarities were observed with *Saccharospirillum impatiens* EL-105<sup>T</sup> (96.46 %), *Saccharospirillum salsuginis* YIM-Y25<sup>T</sup> (96.32 %) and *Saccharospirillum aestuarii* IMCC 4453<sup>T</sup> (95.17 %). The next closest matches were with other genera and below 95.0 %. Phylogenetic analyses revealed that the strain forms a robust clade with other species of the genus *Saccharospirillum*. The main respiratory quinone was Q-8 and the major fatty acids were C<sub>16:0</sub> and summed feature 8 (C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c). The DNA G+C content was 55.2 mol%. Molecular, physiological and biochemical differences between strain CPA1<sup>T</sup> and other type strains of species of the genus *Saccharospirillum* support the addition of this novel species to the genus, and the name *Saccharospirillum correiae* sp. nov. is proposed, with CPA1<sup>T</sup> (= CECT 9131<sup>T</sup> = LMG 29516<sup>T</sup>) as the type strain.

## Keywords

*Saccharospirillaceae*, endophytic, salt marsh, halophytes, taxonomy

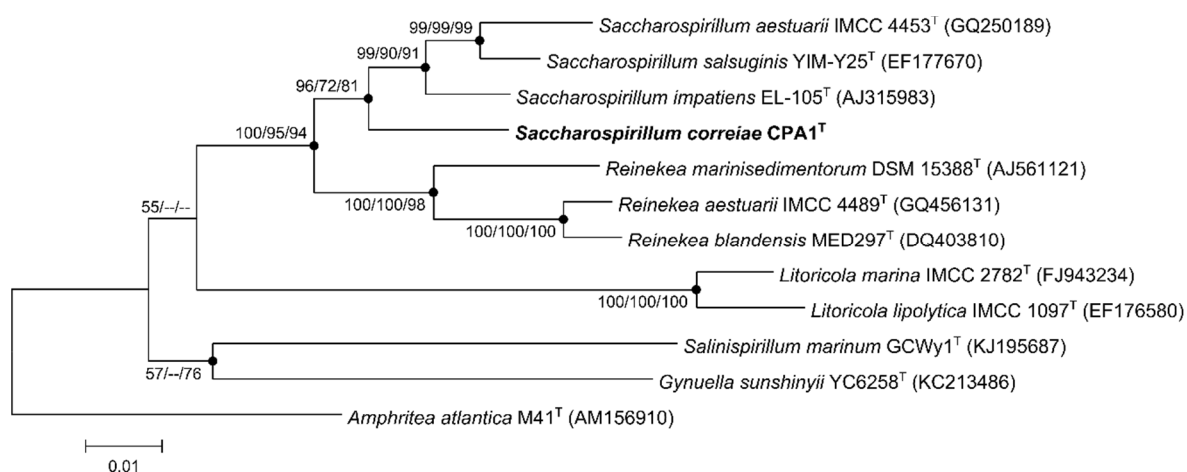
## Main text

The genus *Saccharospirillum*, established by Labrenz et al. (2003), belongs to the family *Saccharospirillaceae* and, at the time of writing, it comprises three species with validly published names, all of which have been isolated from saline environments: *Saccharospirillum impatiens* from a hypersaline lake (Labrenz et al., 2003), *Saccharospirillum salsuginis* from subterranean brine (Chen et al., 2009), and *Saccharospirillum aestuarii* from mudflats (Choi et al., 2011). The genus has been described as Gram-negative spirilla, non-spore forming, containing poly- $\beta$ -hydroxybutyrate, which may form coccoid bodies in older cultures, and positive for peroxidase, catalase and cytochrome oxidase activities, with Q-8 as the prevalent respiratory quinone, and C<sub>16:1</sub>  $\omega$ 7c, C<sub>16:0</sub> and C<sub>18:1</sub>  $\omega$ 7c as the predominant fatty acids (Labrenz et al., 2003). This genus has also been described as comprising obligate aerobes, microaerophiles or facultative anaerobes. The motility and presence of flagella has been determined to be species-dependent (Choi et al., 2011).

In a study of the bacterial diversity associated with the internal tissues of the halophyte *Halimione portulacoides*, 665 isolates were obtained and characterized (Fidalgo et al., 2016). Briefly, healthy specimens of the halophyte were collected in Ria de Aveiro, Portugal, and above- and belowground tissues were separated and surface sterilized. Dilutions of macerated tissues were plated in Tryptic Soy Agar (TSA, Merck, Germany), R2A (Merck, Germany) and Marine Agar (MA, Difco Laboratories, France). The present study focuses on strain CPA1<sup>T</sup>, an isolate obtained from surface sterilized aboveground tissues of *H. portulacoides*. Strain CPA1<sup>T</sup> was routinely streaked on MA medium and incubated under aerobic conditions at 28 °C.

Genomic DNA of the strain was extracted using a Genomic DNA Purification kit #0513 (Thermo Scientific, USA), following the manufacturer's instructions. The 16S rRNA gene was amplified by PCR with universal primers 27F and 1492R, as described elsewhere (Fidalgo et al., 2016), and sequenced with primers 27F (Lane, 1991) and 704F (Kaksonen et al., 2006). Analysis of similarity using a near full length 16S rRNA gene sequence (1418 nt) also gave evidence that strain CPA1<sup>T</sup> belongs to the genus *Saccharospirillum*, as highest similarities were seen with this genus: *S. impatiens* EL-105<sup>T</sup> (96.46 %), *S. salsuginis* YIM-Y25<sup>T</sup> (96.32 %) and *S. aestuarii* IMCC 4453<sup>T</sup> (95.17 %). The next closest matches occurred with type strains of the genus *Reinekea*, with less than 95.0 % 16S rRNA gene sequence similarities. Similarities with type strains belonging to other genera were below 92.0 %.

For phylogenetic analyses, the sequences of closely related taxa were obtained from the EzTaxon database (Kim et al., 2012). Sequence alignments were carried out with Clustal Omega (McWilliam et al., 2013), and BioEdit version 7.2.5 (Hall, 1999) was used to edit the aligned sequences. Phylogenetic analyses were performed using MEGA version 6.0 (Tamura et al., 2013) by using Kimura's two-parameter model (Kimura, 1980) in the reconstruction of neighbour-joining (NJ, Saitou & Nei, 1987), maximum-likelihood (ML, Felsenstein, 1981) and maximum parsimony (MP, Fitch, 1971) trees. Bootstrap values based on 1000 replications were obtained in these phylogenetic analyses. The topology of the ML tree was the same as the MP tree. Figure 4.3 shows that strain CPA1<sup>T</sup> forms a robust clade with the three species of the genus *Saccharospirillum* with validly published names. High bootstrap values were observed, ranging from 72 to 96 %, depending on the tree building method, confirming CPA1<sup>T</sup> as a potential novel member of this genus.



**Figure 4.3** Neighbour-joining (NJ) tree showing the phylogenetic positions of strain CPA1<sup>T</sup> and representatives of other related taxa, based on 16S rRNA gene sequences. Filled circles indicate nodes that were also recovered in the maximum-parsimony (MP) tree and the maximum-likelihood (ML) tree, based on the same sequences. Bootstrap values (expressed as percentages of 1000 replications)  $\geq 50$  % are shown at branching points for NJ/ML/MP trees. *Amphritea atlantica* M41<sup>T</sup> was used as an outgroup. Accession numbers for the type strains are shown in parenthesis. Bar, 0.01 nt substitution rate ( $K_{nuc}$ ) units.

Optimal growth conditions for strain CPA1<sup>T</sup> were determined using MA medium. The temperature range for growth and the optimum were tested by incubating at 4, 18, 26, 30, 37, 42.5 and 50 °C. At the optimum growth temperature, the pH range and optimum were assayed using pH values from 4 to 12, at intervals of 0.5. At the optimum growth temperature and pH, the strain's growth

requirements for, and tolerance to salt (NaCl) was assessed using the following NaCl concentrations: 0.5 (TSA without additional NaCl), 2 (MA without additional NaCl), 3 [MA supplemented with 1 % NaCl (w/v)], 5, 10, 12, 15 and 20 % NaCl (w/v).

Growth in biochemical and phenotypic tests was determined on MA, and incubation was conducted at 30 °C for 48 h, unless otherwise stated. A Gram staining kit (Merck, Germany) was used according to the manufacturer's instructions. Determination of cell size, as well as morphology and motility was carried out using light microscopy (Nikon 80i). The hanging-drop method (Bowman, 2000) was used to assess gliding motility, after growing the strain in half strength MA for 72 h, and using cavity slides. Catalase and oxidase activities were determined using H<sub>2</sub>O<sub>2</sub>-reagent (Liofilchem, Italy) and oxidase strips (Liofilchem, Italy), respectively. Growth in thioglycollate medium (Merck, Germany) was used to assess oxygen metabolism, and incubation was followed for 7 days. Production of H<sub>2</sub>S was assessed using Kligler's Iron Agar (Merck, Germany). Ability to hydrolyse starch, casein, cellulose and xylan was assessed by methods described in Fidalgo et al. (2016).

Biochemical tests were carried out using API 20NE, API ZYM and API 50CH strips (bioMérieux, France) according to the manufacturer's instructions, except for using 0.9 % (w/v) NaCl solution for preparing cell suspensions. *Saccharospirillum impatiens* DSM 12546<sup>T</sup> was tested under the same conditions for comparison. These results are given in the species description and Table 4.5. Some results of the API 50CH tests of *S. impatiens* DSM 12546<sup>T</sup> differed from the original description by Labrenz et al. (2003). The differences observed may result from the use of different suspension medium and different incubation temperatures.

Determination of G+C content was performed by HPLC (Mesbah et al., 1989). The result obtained for strain CPA1<sup>T</sup> was 55.2 %, which is close to the values determined for other species of the genus *Saccharospirillum*, which are 54.5 to 54.8 mol% for *S. impatiens* EL-105<sup>T</sup> (Labrenz et al., 2003), 58.5 mol% for *S. salsuginis* YIM-Y25<sup>T</sup> (Chen et al., 2009), and 56.5 mol% for *S. aestuarii* IMCC 4453<sup>T</sup> (Choi et al., 2011).

**Table 4.5** Differential characteristics of strain CPA1<sup>T</sup> and species of the genus *Saccharospirillum* with validly published names.

Characteristic	CPA1 <sup>T</sup>	<i>S. impatiens</i> DSM 12546 <sup>T</sup> ‡	<i>S. salsuginis</i> YIM-Y25 <sup>T</sup> §	<i>S. aestuarii</i> IMCC 4453 <sup>T</sup>
Cell morphology	Spirilla	Spirilla <sup>a</sup>	Spirilla	Curved rods
Motility	Motile	Motile <sup>a</sup>	Motile	Non motile
Relation to oxygen	Obligately aerobic	Aerobic to microaerophilic <sup>a</sup>	Obligately aerobic	Facultatively anaerobic
Temperature range	18–42.5 °C	< 2.5–43 °C <sup>a</sup>	15–50 °C	10–42 °C
pH range	5–11.5	5.5–9.5 <sup>a</sup>	6–10	5–12
NaCl (w/v) range	0.5–5 %	< 1–15 % <sup>a</sup> †	1–15 %	0.5–10 %
H <sub>2</sub> S production	–	+ <sup>a</sup>	–	–
Hydrolysis of:				
Casein	+	nd	–	+
Starch	–	variable <sup>a</sup>	–	–
Carboxymethyl cellulose	–	nd	nd	+
API 20NE results:				
Hydrolysis of gelatin	+	– *	–	+
Reduction of nitrates to nitrites	–	– *	+	+
Indole production	–	–	+	–
Assimilation of D-Mannose and D-Maltose	–	–	+	nd
Assimilation of Malic acid	–	– *	+	nd
API ZYM results:				
Alkaline phosphatase, Esterase (C4), Esterase lipase (C8), α-Glucosidase	+	w	+	nd
Naphthol-AS-BI- phosphohydrolase, β-Galactosidase	w	w	+	nd
Leucine arylamidase	+	+ *	+	nd
Acid phosphatase	+	–	+	nd
N-acetyl-β-glucosaminidase	+	–	nd	nd
α-fucosidase	w	–	nd	nd
API 50CH results:				
D-Ribose, D-Galactose, D-Glucose, D-Fructose, D-Mannose, D-Cellobiose, D-Maltose, D-Trehalose, Glycogen, L-Rhamnose, D-Lactose (bovine), D-Saccharose (sucrose)	–	– *	+	nd

Inositol, D-Manitol, D-Sorbitol, L-Arabinose, D-Raffinose	–	–	+	nd
Glycerol, Salicin	–	– *	–	nd
Aesculin ferric citrate, D-Melibiose, Starch, D-Turanose, D-Fucose, Gentiobiose	–	– *	nd	nd
DNA G+C content (mol%)	55.2	54.5 - 54.8 <sup>a</sup>	58.5	56.5

+, Positive; w, weakly positive; –, negative; nd, not determined; <sup>a</sup>, data from Labrenz et al. (2003); †, artificial sea water; \*, result differed from that published in Labrenz et al. (2003); ‡, data from this study unless otherwise stated; §, data from Chen et al. (2009); ||, data from Choi et al. (2011).

Assessment of fatty acid profiles, polar lipids and quinones was conducted as described by Proença et al. (2014) and performed in parallel with similar assessments of *S. impatiens* DSM 12546<sup>T</sup>. To obtain cells for fatty acid profile determination, growth was obtained on MA at 30 °C for 48 h. The main fatty acids observed for strain CPA1<sup>T</sup> were C<sub>16:0</sub> and summed feature 8 (C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c), which add up to over 78 % of the total (complete profile in Table 4.6). Summed features arise when the equivalent chain length value obtained corresponds to a fatty acid that cannot be separated from another fatty acid. Consequently, the relative concentration of the possible two or more fatty acids is given as a singular value (da Costa et al., 2011). The results are in accordance with what has previously been described for the genus. Growth for determination of quinones and polar lipids was obtained using Marine Broth (MB) and incubating at 30 °C for 72 h. The main quinone detected was Q-8, which is in accordance with data for the genus. Q-9 was also detected, but in minor amounts. The polar lipid profiles obtained for strain CPA1<sup>T</sup> and *S. impatiens* DSM 12546<sup>T</sup> were very similar, suggesting that both strains belong to the same genus (Supplementary Figure S4.3). The lipids phosphatidylethanolamine, monomethylphosphatidylethanolamine and phosphatidylcholine were identified in the profile of strain CPA1<sup>T</sup>, but diphosphatidylglycerol and phosphatidylglycerol were not detected.

**Table 4.6** Fatty acid composition of CPA1<sup>T</sup> and type strains of species of the genus *Saccharospirillum*.

Fatty acid	CPA1 <sup>T</sup>	<i>S. impatiens</i> DSM 12546 <sup>T</sup> *	<i>S. impatiens</i> EL-105 <sup>T</sup> †	<i>S. salsuginis</i> YIM-Y25 <sup>T</sup> ‡	<i>S. aestuarii</i> IMCC 4453 <sup>T</sup> §
Unknown ECL					
15.272	tr	2.2	–	–	–
17.314	1.7	–	–	–	–
Saturated					
C <sub>12:0</sub> aldehyde	–	–	–	–	4.2
C <sub>14:0</sub>	1.5	tr	–	tr	tr
C <sub>16:0</sub>	31.4	27.4	19.0	11.4	24.3
C <sub>17:0</sub>	tr	tr	–	1.2	tr
C <sub>18:0</sub>	tr	tr	–	–	1.1
Unsaturated					
C <sub>16:1</sub> ω7c	–	–	21.8	–	–
C <sub>17:1</sub> ω6c	–	tr	–	1.2	tr
C <sub>17:1</sub> ω8c	tr	tr	–	2.3	tr
C <sub>18:1</sub> ω7c	–	–	51.2	53.4	–
C <sub>19:1</sub>	–	–	1.9	–	–
Branched-chain					
iso-C <sub>14:0</sub> 3-OH	–	–	–	1.4	–
iso-C <sub>16:0</sub>	tr	tr	–	13.3	7.9
iso-C <sub>18:0</sub>	–	–	–	tr	1.5
iso-C <sub>18:1</sub> H	–	–	–	1.2	–
Hydroxyl					
C <sub>14:0</sub> 3-OH	–	–	1.6	–	–
C <sub>14:1</sub> 3-OH	–	–	2.3	–	–
Cyclic					
cyclo-C <sub>17:0</sub>	4.0	tr	–	–	tr
cyclo-C <sub>19:0</sub> ω8c	2.1	–	–	1.4	tr
Summed feature					
2	5.4	2.8	–	3.8	–
3	3.3	15.9	–	–	–
8	46.8	44.5	–	–	46.6
3 <sup>a</sup>	–	–	–	5.2	6.7
7 <sup>b</sup>	–	–	–	–	1.4

Values represent percentage of total fatty acids. ECL, equivalent chain length; –, not detected; tr, trace amount (< 1 %). Summed features represent groups of two or more fatty acids that could not be separated. Summed feature 2 contains C<sub>12:0</sub> aldehyde, C<sub>14:0</sub> 3-OH and/or iso-C<sub>16:1</sub> and/or Unknown ECL 10.927; Summed feature 3 contains C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH; Summed feature 8 contains C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c; Summed feature 3<sup>a</sup> (from Chen et al., 2009) contains C<sub>16:1</sub> ω7c e/ou C<sub>16:1</sub> ω6c; Summed feature 7<sup>b</sup> (from Choi et al., 2011) contains C<sub>19:1</sub> ω6c e/ou C<sub>19:1</sub> ω7c. \*, data from this study; †, data from Labrenz et al. (2003); ‡, data from Chen et al. (2009); §, data from Choi et al. (2011).

The physiological and biochemical test results for strain CPA1<sup>T</sup> are given in the species description. The phylogenetic and 16S rRNA gene sequencing data, as well as the similarities in biochemical and

physiological characteristics, indicate that strain CPA1<sup>T</sup> belongs to the genus *Saccharospirillum*. This novel strain is distinguishable from other species in the genus with validly published names, and these differentiating characteristics are listed in Table 4.5. As the threshold for genomic delineation of a novel species (97 % 16S rRNA gene sequence similarity; Stackebrandt & Goebel, 1994; Wayne et al., 1987) was not surpassed, there was no need for DNA-DNA relatedness tests. So, CPA1<sup>T</sup> is suggested to represent a novel species of the genus *Saccharospirillum*, and the name *Saccharospirillum correiae* sp. nov. is proposed.

### **Description of *Saccharospirillum correiae* sp. nov.**

*Saccharospirillum correiae* (cor.rei'ae. N.L. gen. masc. n. *correiae* of Correia, in honour of Portuguese microbiologist António Correia).

Cells are Gram-stain-negative, oxidase and catalase positive, motile (not by gliding), obligately aerobic, non-pigmented spirilla (0.47–0.93 µm × 2.33–7.6 µm). Colonies are whitish, opaque in the centre and less so around the regular smooth edges and 0.5–1.5 mm in diameter after incubation in MA for 48 h at 30 °C. Moderately halophilic, mesophilic, and facultatively alkaliphilic, growing at salinities of 0.5 to 5 % NaCl (w/v) with the optimum at 2 % NaCl (w/v). Grows from 18 °C to 42.5 °C (optimum 30 °C) and from pH 5.0 to 11.5 (optimum pH 7.0–7.5). Does not produce H<sub>2</sub>S. Hydrolyses casein and xylan, does not hydrolyse starch and cellulose. In API 20NE strips, it is positive for β-glucosidase, protease (hydrolysis of gelatine) and β-galactosidase (hydrolysis of para-nitrophenyl-β-D-galactopyranose); and negative for the reduction of nitrates to nitrites, nitrates to nitrogen, indole production, fermentation (D-glucose), arginine dihydrolase, urease, assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetylglucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid. In API ZYM strips, it is positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase (weakly), β-galactosidase (weakly), α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase and α-fucosidase (weakly); and negative for lipase (C14), cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-glucuronidase and α-mannosidase. In API 50CH strips it does not produce acid from glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-β-D-xylopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-manitol, D-sorbitol, methyl-α-D-mannopyranoside, methyl-α-D-glucopyranoside, N-acetylglucosamine, amygdalin, arbutin, aesculin ferric citrate, salicin, D-celiobiose, D-maltose, D-



lactose (bovine), D-melibiose, D-saccharose (sucrose), D-trehalose, inulin, D-melezitose, D-raffinose, amidon (starch), glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate. The main fatty acids are C<sub>16:0</sub> and summed feature 8 (C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c), and the main respiratory quinone is Q-8. The main polar lipids comprise phosphatidylethanolamine, monomethylphosphatidylethanolamine and unidentified polar lipids.

The type strain, CPA1<sup>T</sup> (=CECT 9131<sup>T</sup> = LMG 29516<sup>T</sup>) was isolated from the surface sterilized aboveground tissues of the halophyte *Halimione portulacoides*. The G+C content of the DNA of the type strain is 55.2 mol%.

## Acknowledgements

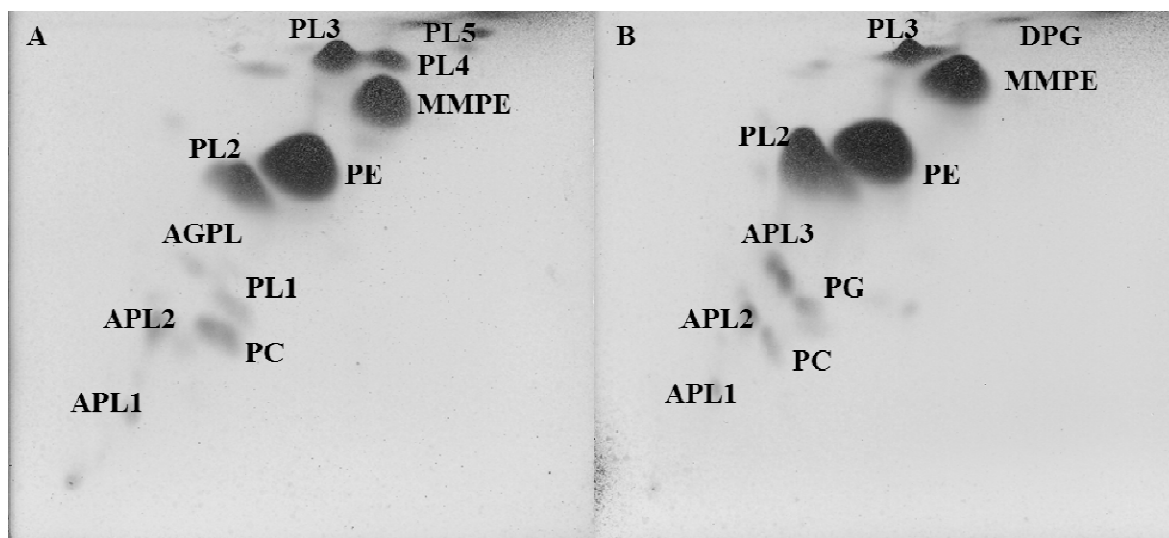
This work was supported by European Funds (FEDER) through COMPETE and by National Funds through the Portuguese Foundation for Science and Technology (FCT) within project PhytoMarsh (PTDC/AAC-AMB/118873/2010 – FCOMP-01-0124-FEDER-019328). FCT is acknowledged for financing to CESAM (UID/AMB/50017/2013 – POCI-01-0145-FEDER-007638), Artur Alves (FCT Investigator Programme – IF/00835/2013), Isabel Henriques (FCT Investigator Programme – IF/00492/2013), Diogo Neves Proença (postdoctoral grant SFRH/BPD/100721/2014) and Cátia Fidalgo (PhD grant – SFRH/BD/85423/2012).

## References

- Bowman, J. P. (2000).** Description of *Cellulophaga algicola* sp. nov., isolated from the surfaces of Antarctic algae, and reclassification of *Cytophaga uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Cellulophaga uliginosa* comb. nov. *Int J Syst Evol Microbiol* **50**, 1861–1868.
- Chen, Y.-G., Cui, X.-L., Li, Q.-Y., Wang, Y.-X., Tang, S.-K., Liu, Z.-X., Wen, M.-L., Peng, Q. & Xu, L.-H. (2009).** *Saccharospirillum salsuginis* sp. nov., a gammaproteobacterium from a subterranean brine. *Int J Syst Evol Microbiol* **59**, 1382–1386.
- Choi, A., Oh, H.-M. & Cho, J.-C. (2011).** *Saccharospirillum aestuarii* sp. nov., isolated from tidal flat sediment, and an emended description of the genus *Saccharospirillum*. *Int J Syst Evol Microbiol* **61**, 487–492.
- da Costa, M. S., Albuquerque, L., Nobre, M. F. & Wait, R. (2011).** The identification of fatty acids in Bacteria. In *Methods in Microbiology (Taxonomy of Prokaryotes)*, vol. 38, pp. 183–196. Edited by F. A. Rainey & A. Oren. Amsterdam: Elsevier.
- Felsenstein, J. (1981).** Evolutionary trees from DNA sequences: a maximum-likelihood approach. *J Mol Evol* **17**, 368–376.
- Fidalgo, C., Henriques, I., Rocha, J., Tacão, M. & Alves, A. (2016).** Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization, and influence of metal(loid) contamination. *Environ Sci Pollut Res* **23**(10), 10200–10214.
- Fitch, W. M. (1971).** Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* **20**, 406–416.
- Hall, T. A. (1999).** BioEdit: a user-friendly biological sequence alignment edit and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* **41**, 95–98.

- Kaksonen, A. H., Plumb, J. J., Robertson, W. J., Spring, S., Schumann, P., Franzmann, P. D. & Puhakka, J. A. (2006).** Novel thermophilic sulfate-reducing bacteria from a geothermally active underground mine in Japan. *Appl Environ Microbiol* **72**(5), 3759-3762.
- Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H., Yi, H., Won, S. & Chun, J. (2012).** Introducing EzTaxon: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* **62**, 716–721.
- Kimura, M. (1980).** A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Labrenz, M., Lawson, P. A., Tindall, B. J., Collins, M. D. & Hirsch, P. (2003).** *Saccharospirillum impatiens* gen. nov., sp. nov., a novel  $\gamma$ -Proteobacterium isolated from hypersaline Ekho Lake (East Antarctica). *Int J Syst Evol Microbiol* **53**, 653-660.
- Lane, D. J. (1991).** 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*, pp. 115–175. Edited by E. Stackebrandt & M. Goodfellow. New York, NY: John Wiley and Sons.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013).** Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **41**, (Web Server issue):W597-600 doi:10.1093/nar/gkt376.
- Mesbah, M., Premachandran, U. & Whitman, W.B. (1989).** Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* **39**, 159–167.
- Proença, D. N., Nobre, M. F. & Morais, P. V. (2014).** *Chitinophaga costaii* sp. nov., an endophyte of *Pinus pinaster*, and emended description of *Chitinophaga niabensis*. *Int J Syst Evol Microbiol* **64**, 1237-1243.
- Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Stackebrandt, E. & Goebel, B. M. (1994).** Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* **44**, 846-849.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A. & Kumar, S. (2013).** MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. & Trüper, H. G. (1987).** Report of the Ad Hoc Committee on reconciliation of approaches to bacterial systematics. *Int J Syst Evol Microbiol* **37**, 463-464.

## Supplementary material



**Supplementary Figure S4.3** Two-dimensional thin-layer chromatography of the total polar lipids of strain CPA1<sup>T</sup> (A) and *Saccharospirillum impatiens* DSM 12546<sup>T</sup> (B) stained with 5% ethanolic molybdophosphoric acid. PE, phosphatidylethanolamine; MMPE, monomethyl-phosphatidylethanolamine; PC, phosphatidylcholine; PG, phosphatidylglycerol; DPG, diphosphatidylglycerol; APL1-3, unidentified aminophospholipids; PL1-5, unidentified phospholipids; AGPL, unidentified aminoglycophospholipid.

### **4.3 *Altererythrobacter halimionae* sp. nov. and *Altererythrobacter endophyticus* sp. nov., two endophytes from the salt marsh plant *Halimione portulacoides***

#### **Authors**

Cátia Fidalgo<sup>1</sup>, Jaqueline Rocha<sup>1</sup>, Ricardo Martins<sup>1</sup>, Diogo Neves Proença<sup>2</sup>, Paula V. Morais<sup>2,3</sup>, Isabel Henriques<sup>1</sup>, Artur Alves<sup>1</sup>

<sup>1</sup> Departamento de Biologia, CESAM, Universidade de Aveiro, Aveiro, Portugal

<sup>2</sup> CEMUC, University of Coimbra, 3030-788 Coimbra, Portugal

<sup>3</sup> Department of Life Sciences, FCTUC, University of Coimbra, 3000-456 Coimbra, Portugal

#### **Publication status**

In press

Int J Syst Evol Microbiol

doi: 10.1099/ijsem.0.002079

## Abstract

Two Gram-negative, rod-shaped, motile bacterial strains, named CPA5<sup>T</sup> and BR75<sup>T</sup> were isolated from the halophyte *Halimione portulacoides*. Both presented optimum growth at 30 °C, pH 7.0–7.5 and 1–2 % NaCl (w/v) for strain CPA5<sup>T</sup>, and pH 7.5–8.0 and 2 % NaCl (w/v) for strain BR75<sup>T</sup>. Phylogenetic analyses based on 16S rRNA gene sequences affiliated both strains to the genus *Altererythrobacter*. CPA5<sup>T</sup> presented highest 16S rRNA gene sequence similarity with *Altererythrobacter aestuarii* KYW147<sup>T</sup> (96.5 %), followed by *Altererythrobacter namhicola* KYW48<sup>T</sup> (95.9 %), *Novosphingobium indicum* H25<sup>T</sup> (95.6 %), and *Altererythrobacter oceanensis* Y2<sup>T</sup> (95.5 %). BR75<sup>T</sup> displayed highest similarity with *Altererythrobacter marensis* MSW-14<sup>T</sup> (96.5 %), followed by *Altererythrobacter xinjiangensis* S3-63<sup>T</sup>, *Altererythrobacter luteolus* SW-109<sup>T</sup> and *Altererythrobacter indicus* MSSRF26<sup>T</sup> (96.1 %). Neither strain contained Bacteriochlorophyll *a*. The main fatty acids observed for CPA5<sup>T</sup> were C<sub>17:1</sub> ω6c and summed features 3 (C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH) and 8 (C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c). The latter summed feature was the dominant fatty acid observed for strain BR75<sup>T</sup> as well. The major polar lipids were phosphatidylethanolamine, unidentified phospholipids and unidentified glycolipids for both strains. The predominant ubiquinone was Q-10 for both strains, and the DNA G+C content was 63.4 mol% and 58.3 mol% for CPA5<sup>T</sup> and BR75<sup>T</sup>, respectively. Based on phenotypic and genotypic results, both strains represent novel species belonging to the genus *Altererythrobacter* for which the names *Altererythrobacter halimionae* sp. nov. (type strain CPA5<sup>T</sup> = CECT 9130<sup>T</sup>=LMG 29519<sup>T</sup>) and *Altererythrobacter endophyticus* sp. nov (type strain BR75<sup>T</sup> = CECT 9129<sup>T</sup>=LMG 29518<sup>T</sup>) are proposed.

## Keywords

*Erythrobacteraceae*, endophytic, halophytes, taxonomy

## Main text

The genus *Altererythrobacter* was described in 2007 (Kwon et al., 2007), emended in 2012 (Xue et al., 2012), and 2016 (Xue et al., 2016), and belongs to the family *Erythrobacteraceae* (Lee et al., 2005). At the time of writing the genus contains 22 validly published species, several of which frequently isolated from marine and estuarine environments (e.g. Lai et al., 2009; Seo & Lee, 2010; Park et al., 2011; Park et al., 2016). Its occurrence in association with plants is rare and there is only one species that has been isolated from the rhizosphere of wild rice (Kumar et al., 2008).

The genus *Altererythrobacter* comprises Gram-negative bacteria that do not produce H<sub>2</sub>S. Cells cannot grow in anaerobic conditions and nitrate is not reduced. Cell suspensions and colonies are yellow, and the methanol-soluble pigment indicates absence of Bacteriochlorophyll *a* (BChl *a*). The main quinone is Q-10 (Kwon et al., 2007) and the major polar lipids are phosphatidylethanolamine, phosphatidylglycerol, diphosphatidylglycerol and sphingoglycolipid (Xue et al., 2012). The DNA G+C content range is 54.5–67.5 mol%, and catalase reaction can be positive or negative (Xue et al., 2016). The major fatty acids include C<sub>18:1</sub> ω7c (Kwon et al., 2007), C<sub>16:1</sub> ω7c and C<sub>17:1</sub> ω6c.

The diversity of the endophytic community of the halophyte *Halimione portulacoides* was assessed in a salt marsh in Aveiro, Portugal. Briefly, healthy specimens of the halophyte were collected, aboveground and belowground tissues from these specimens were separated, surface-sterilized, macerated in phosphate buffer solution and studied for their bacterial diversity (Fidalgo et al., 2016). This study focuses on two strains obtained in those isolation efforts: strain CPA5<sup>T</sup>, isolated from the aboveground tissues; and BR75<sup>T</sup>, isolated from the belowground tissues of the halophyte. Strains CPA5<sup>T</sup> and BR75<sup>T</sup> were originally isolated from and routinely cultured on Marine Agar (MA, Difco Laboratories, France) culture medium, at 28 °C, under aerobic conditions.

Genomic DNA was extracted, subjected to PCR amplification for 16S rRNA gene, and sequenced as described elsewhere (Fidalgo et al., 2016). The primers 27F (Lane, 1991) and 704F (Kaksonen et al., 2006) were used for sequencing the 16S rRNA gene. The nearly full length sequences obtained for CPA5<sup>T</sup> (1412 nt) and BR75<sup>T</sup> (1406 nt) were used for similarity analyses using the Identify tool present in the EzTaxon platform (Kim et al., 2012). For strain CPA5<sup>T</sup>, the closest matches were observed with type strains *Altererythrobacter aestuarii* KYW147<sup>T</sup> (96.5 % similarity of the 16S rRNA gene sequence), followed by *Altererythrobacter namhicola* KYW48<sup>T</sup> (95.9 %), *Novosphingobium indicum* H25<sup>T</sup> (95.6 %), and *Altererythrobacter oceanensis* Y2<sup>T</sup> (95.5 %). For strain BR75<sup>T</sup>, the most closely related type strains were *Altererythrobacter marensis* MSW-14<sup>T</sup> (96.5 %), followed by *Altererythrobacter xinjiangensis* S3-63<sup>T</sup> and *Altererythrobacter luteolus* SW-109<sup>T</sup> (96.1 %) and

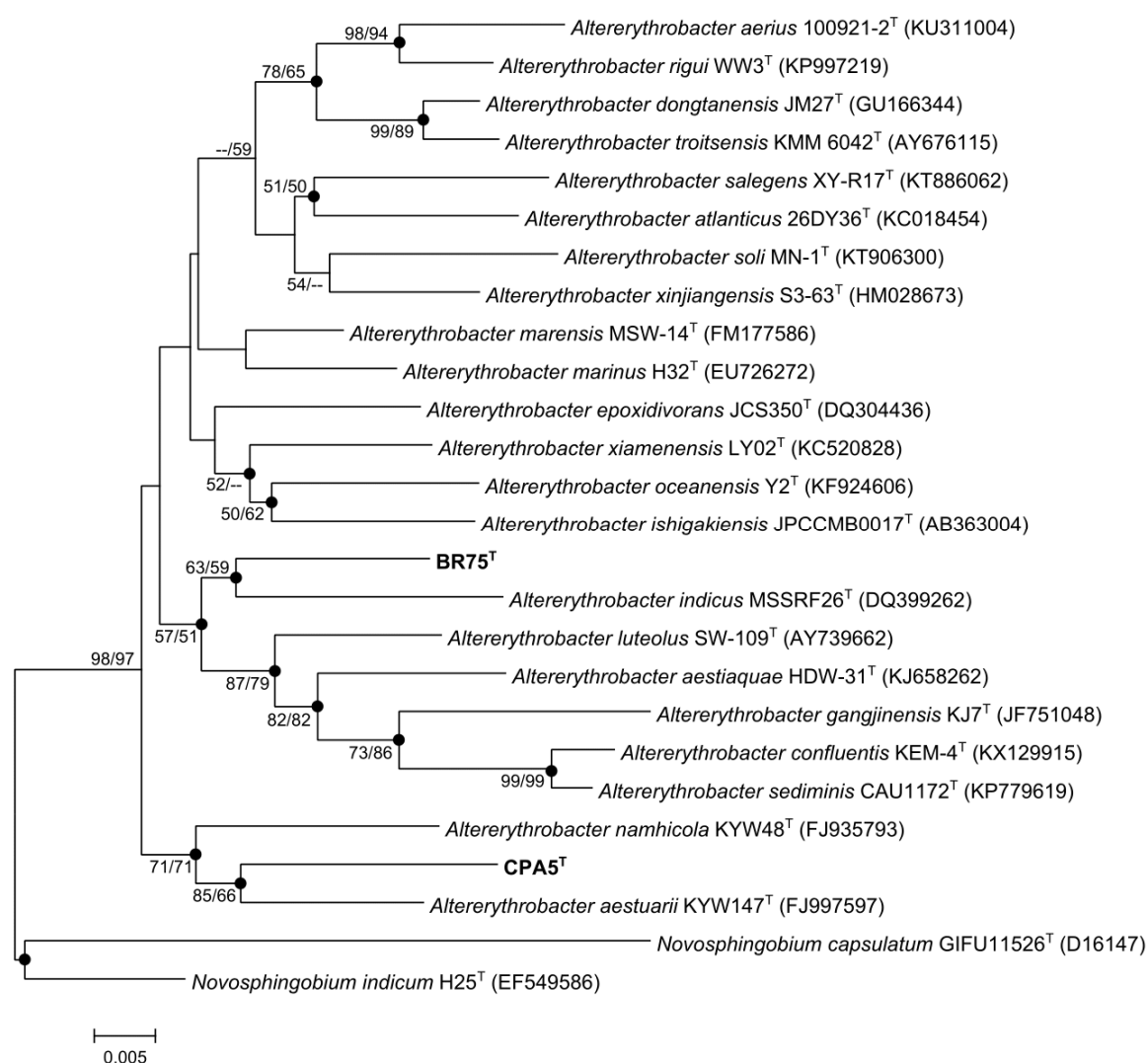
*Altererythrobacter indicus* MSSRF26<sup>T</sup> (96.1 %). 16S rRNA gene sequence similarity percentages to other type strains were below 95.5 % and 96.0 % for CPA5<sup>T</sup> and BR75<sup>T</sup>, respectively.

The 16S rRNA gene sequences of strains CPA5<sup>T</sup> and BR75<sup>T</sup> were aligned with sequences of related type strains, retrieved from the EzTaxon database (Kim et al., 2012). The sequences were then aligned using Clustal Omega (McWilliam et al., 2013) and edited using BioEdit version 7.2.5 (Hall, 1999). MEGA version 6.0 (Tamura et al., 2013) was used to cluster the sequences applying different methods: neighbour-joining (NJ, Saitou & Nei, 1987) and maximum-likelihood (ML, Felsenstein, 1981). The Kimura two-parameter model (Kimura, 1980) was used in clustering, and bootstrap values based on 1000 replications were obtained in these phylogenetic analyses. The obtained phylogenetic trees clearly placed both strains in independent clusters in the genus *Altererythrobacter* (Figure 4.4). A more extended overview of the placement of strains CPA5<sup>T</sup> and BR75<sup>T</sup> in the context of the family *Erythrobacteraceae* is represented in Supplementary Figure S4.4.

The optimal conditions for growth were tested using a base of MA medium. The range and optimum conditions were tested for temperature first, then pH and finally for NaCl concentration. Tests were performed by incubating strains at 4, 18, 26, 30, 37, 42.5 and 50 °C, pH from 4 to 12 in 0.5 intervals, and NaCl tolerance was tested using concentrations of 0, 0.5, 1, 2, 3, 5, 10, 15 and 20 % NaCl (w/v) in a medium composed of 5 g L<sup>-1</sup> yeast extract (Alfa Aesar, Massachusetts, USA) and 10 g L<sup>-1</sup> tryptone casein peptone (Amresco, Texas, USA). Optimum temperature for growth was observed at 30 °C for both strains. Optimum growth for CPA5<sup>T</sup> was observed at pH 7-7.5 and 1–2 % NaCl (w/v), and for strain BR75<sup>T</sup> at pH 7.5-8 and 2 % NaCl (w/v).

Biochemical and phenotypic tests were performed with cells grown on MA medium for 48 h, at 30 °C. Gram staining reaction was performed with a kit, and manufacturer's instructions (Merck, Germany) were followed. Catalase and oxidase activities were assessed using the H<sub>2</sub>O<sub>2</sub>-reagent and oxidase strips, respectively (both from Liofilchem, Italy). Light microscopy was used for determination of cell size, morphology and motility. Additionally, cells grown in half-strength MA for 72 h were placed on cavity slides and gliding motility was assessed by hang-drop method (Bowman, 2000). Oxygen metabolism was assessed by observing growth on thioglycollate medium (Merck, Germany) for 7 days. The ability to produce H<sub>2</sub>S was assessed using Kligler's Iron Agar (Merck, Germany). Ability to hydrolyse starch, Tween 20, xylan, casein and cellulose, and to produce indole-3-acetic acid (IAA) were assessed as described in Fidalgo et al. (2016).





**Figure 4.4** Neighbour-joining (NJ) tree showing the phylogenetic positions of strains CPA5<sup>T</sup> and BR75<sup>T</sup> and representatives of other related taxa based on 16S rRNA gene sequences. Full circles denote nodes that were also recovered in the maximum-likelihood (ML) tree, based on the same sequences. Bootstrap values (expressed as percentages of 1000 replications) which are  $\geq 50\%$  are shown at branch points, following the order: NJ/ML. Accession numbers for the type strains are shown in parenthesis. Bar, 0.005 nt substitution rate ( $K_{nuc}$ ) units.

To assess presence of Bacteriochlorophyll *a* and absorbance peaks of the pigments, cells were grown on Marine Broth (MB), washed once with distilled water, vigorously resuspended in 90 % (v/v) acetone and centrifuged. The supernatant was then removed and kept at 4 °C in the dark overnight. Absorption peaks were assessed from 300 to 800 nm using the Thermo Spectroscopy Genesys 6.

Additional biochemical tests were performed for strains CPA5<sup>T</sup>, BR75<sup>T</sup> as well as type strains *Altererythrobacter marensis* KCTC 22370<sup>T</sup> and *Altererythrobacter aestuarii* KCTC 22735<sup>T</sup>, using API 20NE, API ZYM and API 50CH strips (bioMérieux, France) following the manufacturer's instructions, except for using 0.9 % (w/v) NaCl to prepare inocula. The results for biochemical and phenotypic tests are detailed in the description of the new species, and differentiating characteristics are stated in Table 4.7.

The assessment of respiratory quinones, polar lipids and profile for fatty acids was conducted as described in Proença et al. (2014) and performed with strains CPA5<sup>T</sup>, *Altererythrobacter aestuarii* KCTC 22735<sup>T</sup>, BR75<sup>T</sup> and *Altererythrobacter marensis* KCTC 22370<sup>T</sup> simultaneously. Cells were grown in MB at 30 °C for 48 h to obtain biomass for quinone and polar lipids assays. The main quinone detected for all strains was Q-10, and Q-9 and Q-8 were detected in minor amounts. The main fatty acids observed for CPA5<sup>T</sup> were C<sub>17:1</sub> ω6c (13.8 %) and summed features 3 (C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH; 21.4 %) and 8 (C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c; 32.6 %), comprising over 67 % of total fatty acids. For strain BR75<sup>T</sup> the main fatty acids were comprised in summed feature 8, representing 76.3 % of total fatty acids. The results are in accordance to what is observed in other *Altererythrobacter* species, seen as the characteristic fatty acid of the genus (C<sub>18:1</sub> ω7c) is present in summed feature 8 of our analysis. The complete fatty acid composition for all tested strains is presented in Table 4.8. The polar lipid profiles obtained are depicted in Supplementary Figure S4.5. For strain CPA5<sup>T</sup> the polar lipids detected in major amounts included phosphatidylethanolamine (PE), an unidentified glycolipid (GL2) and three unidentified phospholipids (PL2, PL3 and PL5). The profile for the phylogenetically close relative *A. aestuarii* KCTC 22735<sup>T</sup> was similar to that obtained for strain CPA5<sup>T</sup>, albeit presenting small differences in regards to the minor polar lipids. For strain BR75<sup>T</sup>, the major polar lipids were PE, an unidentified glycolipid (GL2) and four unidentified phospholipids (PL2, PL3, PL4 and PL5). The profile was highly similar to that of *A. marensis* KCTC 22370<sup>T</sup> and only discrepancies in polar lipids in minor amounts were observed. These results further indicate that strains CPA5<sup>T</sup> and BR75<sup>T</sup> belong to the genus *Altererythrobacter* but present slight differences with the most closely related strains. Determination of G+C content was performed by HPLC (Mesbah et al., 1989). The results obtained (63.4 mol% for CPA5<sup>T</sup> and 58.3 mol% BR75<sup>T</sup>) are in agreement with what has been previously observed in the genus (54.5 to 67.5 mol%; Xue et al., 2016).

**Table 4.7** Differential characteristics of strains CPA5<sup>T</sup>, BR75<sup>T</sup>, and related type strains (data from this study unless otherwise stated).

Characteristic	CPA5 <sup>T</sup>	BR75 <sup>T</sup>	<i>A. aestuarii</i> KCTC 22735 <sup>T</sup>	<i>A. marensis</i> KCTC 22370 <sup>T</sup>
Motility	+	+	– <sup>a</sup>	+ <sup>b</sup>
Catalase activity	+	+	+ <sup>a</sup>	w <sup>b</sup>
Hydrolysis of casein	–	+	– <sup>a</sup>	– <sup>b</sup>
NaCl (w/v) range	0–5 %	0–5 %	0–6 % <sup>a</sup>	0–9 % <sup>b</sup>
pH range	5–11.5	5–11.5	5–11 <sup>a</sup>	6.1–11.1 <sup>b</sup>
Temperature range	18–37 °C	18–37 °C	10–40 °C <sup>a</sup>	4–42 °C <sup>b</sup>
API 20NE results:				
Reduction of nitrates to nitrites	–	–	+ †	–
β-galactosidase (para-nitrophenyl-β-D-galactopyranose)	–	+	–	–
Assimilation of malic acid	–	+	–	+
API ZYM results:				
Esterase (C4), esterase lipase (C8), α-chymotrypsin	+	+	+ †	+
Lipase (C14)	+	w	–	–
Valine arylamidase	+	+	w †	+
Cystine arylamidase	w	–	–	w ‡
Trypsin	w	–	+	w ‡
Acid phosphatase	+	+	–	– ‡
Naphthol-AS-BI-phosphohydrolase	+	+	w †	w ‡
β-galactosidase	–	w	–	–
β-glucuronidase	–	+	–	–
α-glucosidase	–	–	+ †	–
β-glucosidase	–	+	+	–
API 50CH results (acid production):				
Aesculin ferric citrate	–	+	– *	+
Salicin	–	–	–	+
DNA G+C content (mol%)	63.4	58.3	67.2 <sup>a</sup>	63.1 <sup>b</sup>

+, Positive; w, weakly positive; –, negative; <sup>a</sup>, data from Seo & Lee, 2010; <sup>b</sup>, data from Park et al., 2011; †, result differed from that published in Seo & Lee, 2010; ‡, result differed from that published in Park et al., 2011; \*, result differed from that published in Seo & Lee, 2010 where it was obtained with a different methodology.

**Table 4.8** Fatty acid composition of strains CPA5<sup>T</sup>, BR75<sup>T</sup> and related type strains.

Fatty acid	CPA5 <sup>T</sup>	BR75 <sup>T</sup>	<i>A. aestuarii</i> KCTC 22735 <sup>T</sup> (this study)	<i>A. aestuarii</i> KYW147 <sup>T</sup> (Seo & Lee, 2010)	<i>A. marensis</i> KCTC 22370 <sup>T</sup> (this study)	<i>A. marensis</i> MSW-14 <sup>T</sup> (Park et al., 2011)
<b>Saturated</b>						
C <sub>15:0</sub>	1.6	–	tr	3.2	–	1.6
C <sub>16:0</sub>	7.1	5.8	9.0	5.8	4.9	13
C <sub>17:0</sub>	4.7	–	tr	–	–	–
C <sub>18:0</sub>	tr	1.1	tr	–	tr	3.7
<b>Unsaturated</b>						
C <sub>16:1</sub> ω5c	tr	–	1.5	–	1.4	1.9
C <sub>17:1</sub> ω6c	13.8	3.1	6.3	19.9	3.4	6.8
C <sub>17:1</sub> ω8c	1.5	tr	1.1	2.1	–	–
C <sub>18:1</sub> ω5c	1.3	tr	1.4	–	2.1	2.2
C <sub>18:1</sub> ω7c	–	–	–	35.2	–	–
C <sub>18:1</sub> ω7c 11-methyl	9.4	–	13.3	–	24.0	–
<b>Hydroxyl</b>						
C <sub>14:0</sub> 2-OH	2.3	8.1	1.0	7.5	2.3	1.3
C <sub>15:0</sub> 2-OH	tr	tr	tr	3.5	–	–
C <sub>16:0</sub> 2-OH	–	tr	3.1	–	1.5	1.4
<b>Summed feature</b>						
3	21.4	2.6	19.4	22.7	8.7	8.8
7	–	–	–	–	–	52.8
8	32.6	76.3	41.1	–	50.4	–

Values represent percentage of total fatty acids. –, not detected; tr, trace amount (< 1 %). Summed feature 3 contains C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH; Summed feature 7 contains C<sub>18:1</sub> ω9c and/or C<sub>18:1</sub> ω12t and/or C<sub>18:1</sub> ω 7c; Summed feature 8 contains C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c.

Detailed results for each strain are given in the respective species description section. Considering the phylogenetic and 16S rRNA gene sequencing data, similarities in physiological and biochemical traits, it is clear that CPA5<sup>T</sup> and BR75<sup>T</sup> belong to the genus *Altererythrobacter*. Given that the 16S rRNA sequence similarities did not surpass the threshold for genomic delimitation of a new species (97 % sequence similarity; Stackebrandt & Goebel, 1994; Wayne et al., 1987), there was no need to perform DNA-DNA relatedness tests. Strains CPA5<sup>T</sup> and BR75<sup>T</sup> are, nevertheless, distinguishable from validly published species of the genus, as they present differences in certain traits (Table 4.7). Differences between the novel species BR75<sup>T</sup> and the closely related reference strain of *A. marensis* include the ability to hydrolyse casein, to assimilate malic acid, to produce acid from salicin, and

activity of  $\beta$ -galactosidase,  $\beta$ -glucosidase and  $\beta$ -glucuronidase. Between CPA5<sup>T</sup> and *A. aestuarii* differences include motility and activity of lipase, acid phosphatase and  $\beta$ -glucosidase. Accordingly, strains CPA5<sup>T</sup> and BR75<sup>T</sup> represent novel species of the genus *Altererythrobacter*, for which the names *Altererythrobacter halimionae* sp. nov. and *Altererythrobacter endophyticus* sp. nov., respectively, are proposed.

### **Description of *Altererythrobacter halimionae* sp. nov.**

*Altererythrobacter halimionae* (ha.li.mi.o'nae. N.L. gen. n. *halimionae* of the marsh plant *Halimione portulacoides*).

Cells are Gram-negative rods (1.59-3.56  $\mu\text{m}$  L  $\times$  0.5-0.96  $\mu\text{m}$  W), aerobic, motile but not by gliding. Colony on MA after incubation at 30 °C for 48 h is yellow, opaque, with smooth edges and a diameter of 0.5 to 1 mm. Growth is observed from 18 to 37 °C (optimum 30 °C), at pH 5.0 to 11.5 (optimum 7.0-7.5) and in the presence of 0.5 to 5.0 % (w/v) NaCl [optimum 1-2 % (w/v) NaCl], being slightly halophilic. Positive for catalase, oxidase, hydrolysis of Tween 20 and xylan and production of IAA (45.5  $\mu\text{g mL}^{-1}$ ). Does not hydrolyse casein, starch, cellulose, and does not produce H<sub>2</sub>S. Bacteriochlorophyll  $\alpha$  is absent and acetone-soluble peaks were observed at 454 and 482 nm. In API 20NE strips, it is positive for hydrolysis of aesculin ( $\beta$ -glucosidase), and negative for reduction of nitrates, indole production, fermentation of D-glucose, arginine dihydrolase, urease, hydrolysis of gelatin (protease), para-nitrophenyl- $\beta$ -D-galactopyranose ( $\beta$ -galactosidase), assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetylglucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate, phenylacetic acid. In API ZYM strip, it is positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine arylamidase, valine arylamidase,  $\alpha$ -chymotrypsin, acid phosphatase and naphthol-AS-BI-phosphohydrolase; weakly positive for cysteine arylamidase and trypsin. In this strip, it is negative for  $\alpha$ -galactosidase,  $\beta$ -galactosidase,  $\beta$ -glucuronidase,  $\alpha$ -glucosidase,  $\beta$ -glucosidase, N-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -mannosidase and  $\alpha$ -fucosidase. In API 50CH it is negative for acid production from glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl- $\beta$ -D-xylopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl- $\alpha$ -D-mannopyranoside, methyl- $\alpha$ -D-glucopyranoside, N-acetylglucosamine, amygdalin, arbutin, aesculin ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose (bovine), D-melibiose, D-saccharose (sucrose), D-trehalose, inulin, D-melezitose, D-raffinose, amidon (starch), glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-

tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate, potassium 5-ketogluconate. The main quinone is Q-10 and the main fatty acids are C<sub>17:1</sub> ω6c and summed features 3 (C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH) and 8 (C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c). The major polar lipids comprise phosphatidylethanolamine and unidentified polar lipids.

The type strain CPA5<sup>T</sup> (=CECT 9130<sup>T</sup>=LMG 29519<sup>T</sup>) was isolated from the surface-sterilized aboveground tissues of the halophyte *Halimione portulacoides*. The G+C content of the DNA of the type strain is 63.4 mol%.

### **Description of *Altererythrobacter endophyticus* sp. nov.**

*Altererythrobacter endophyticus* (en.do.phy'ti.cus. Gr. pref. *endo* within; Gr. n. *phyton* plant; L. neut. suff. -icus adjectival suffix used with the sense of belonging to; N.L. masc. adj. *endophyticus* within plant, endophytic).

Cells are Gram-negative aerobic rods (1.46-3.95 μm L × 0.59-1.41 μm W), motile but not by gliding. After incubation at 30 °C for 48 h on MA, colony is yellow, opaque, with smooth edges and 0.5 to 1.2 mm in diameter. Growth occurs from 18 to 37 °C (optimum 30 °C), at pH 5.0 to 11.5 (optimum 7.5-8.0) and in the presence of 0.5 to 5.0 % (w/v) NaCl [optimum 2 % (w/v) NaCl], being slightly halophilic. Bacteriochlorophyll *a* is absent, and acetone-soluble peaks are observed at 454-455 and 482-483 nm. Cells are catalase and oxidase positive, and hydrolyse casein, Tween 20 and xylan, and produce IAA (90.8 μg mL<sup>-1</sup>). H<sub>2</sub>S is not produced, and starch and cellulose are not hydrolyzed. In API 20NE strip, it is positive for hydrolysis of aesculin (β-glucosidase), para-nitrophenyl-β-D-galactopyranose (β-galactosidase) and assimilation of malic acid. It is negative for reduction of nitrates, indole production, fermentation of D-glucose, arginine dihydrolase, urease, hydrolysis of gelatin (protease), assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetylglucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, trisodium citrate and phenylacetic acid. In the API ZYM strip, it is positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, α-chymotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, β-glucuronidase, and β-glucosidase, and weakly positive for lipase (C14) and β-galactosidase. It is negative for cystine arylamidase, trypsin, α-galactosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase. In API 50CH it is positive for acid production from aesculin ferric citrate, and negative for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-β-D-xylopyranoside, D-galactose, D-glucose, D-fructose, D-

mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl- $\alpha$ -D-mannopyranoside, methyl- $\alpha$ -D-glucopyranoside, *N*-acetylglucosamine, amygdalin, arbutin, salicin, D-cellobiose, D-maltose, D-lactose (bovine), D-melibiose, D-saccharose (sucrose), D-trehalose, inulin, D-melezitose, D-raffinose, amidon (starch), glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate, potassium 5-ketogluconate. The predominant fatty acids are those contained in summed feature 8 ( $C_{18:1}$   $\omega$ 7c and/or  $C_{18:1}$   $\omega$ 6c), and the principal respiratory quinone is Q-10. The major polar lipids comprise phosphatidylethanolamine and unidentified polar lipids.

The type strain BR75<sup>T</sup> (= CECT 9129<sup>T</sup> = LMG 29518<sup>T</sup>) was isolated from the surface-sterilized belowground tissues of the halophyte *Halimione portulacoides*. The G+C content of the DNA of the type strain is 58.3 mol%.

## Acknowledgements

This work was supported by European Funds (FEDER) through COMPETE and by National Funds through the Portuguese Foundation for Science and Technology (FCT) within project PhytoMarsh (PTDC/AAC-AMB/118873/2010 – FCOMP-01-0124-FEDER-019328). FCT is acknowledged for financing to CESAM (UID/AMB/50017/2013), Artur Alves (FCT Investigator Programme – IF/00835/2013), Isabel Henriques (FCT Investigator Programme – IF/00492/2013), Cátia Fidalgo (PhD grant – SFRH/BD/85423/2012) and Diogo Proença (Post Doc grant – SFRH/BPD/100721/2014).

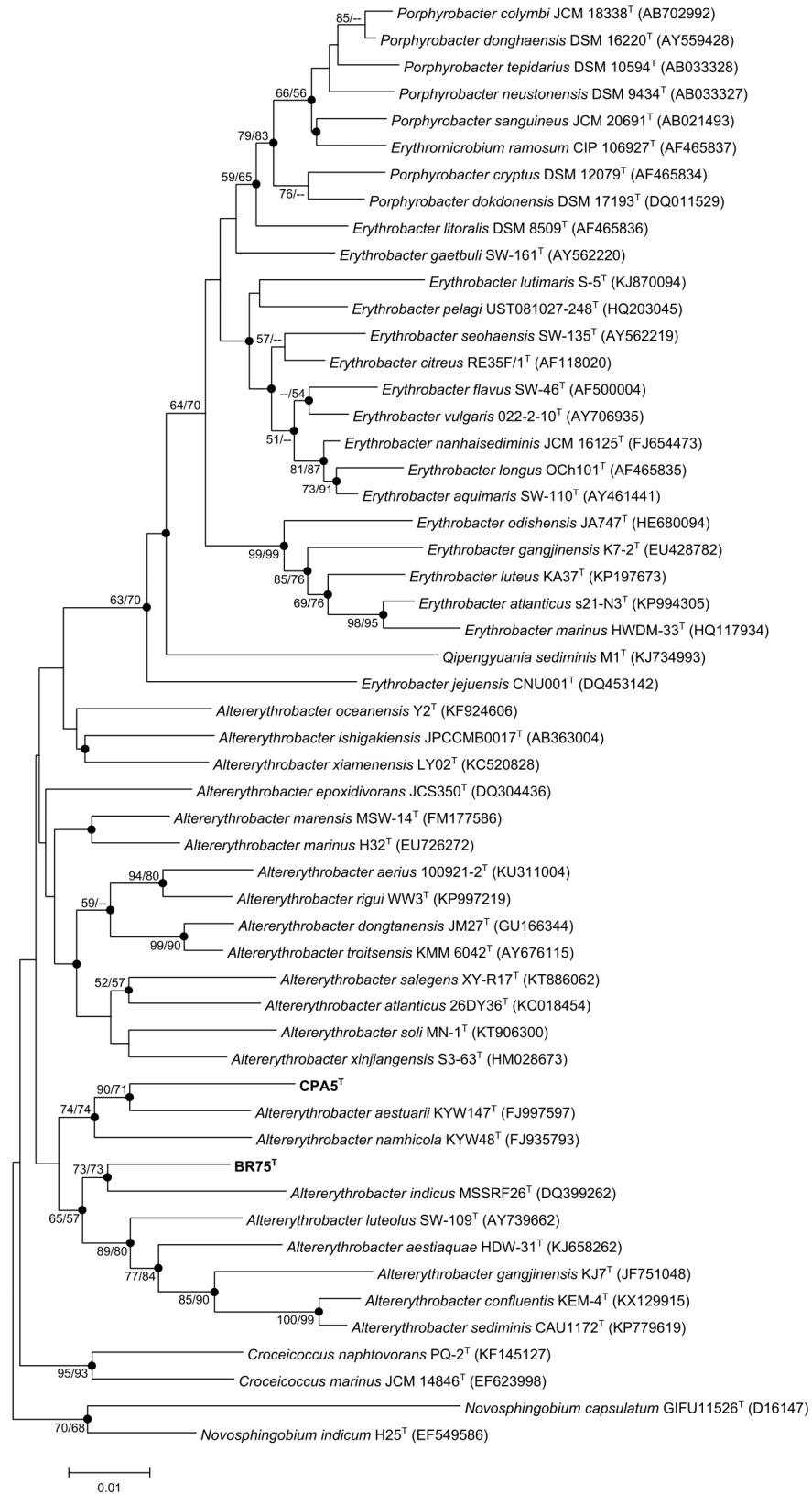
## References

- Bowman, J. P. (2000).** Description of *Cellulophaga algicola* sp. nov., isolated from the surfaces of Antarctic algae, and reclassification of *Cytophaga uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Cellulophaga uliginosa* comb. nov. *Int J Syst Evol Microbiol* **50**, 1861–1868.
- Felsenstein, J. (1981).** Evolutionary trees from DNA sequences: a maximum-likelihood approach. *J Mol Evol* **17**, 368–376.
- Fidalgo, C., Henriques, I., Rocha, J., Tacão, M. & Alves, A. (2016).** Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization, and influence of metal(loid) contamination. *Environ Sci Pollut Res* **23**, 10200–10214.
- Hall, T. A. (1999).** BioEdit: a user-friendly biological sequence alignment edit and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **41**, 95–98.
- Kaksonen, A. H., Plumb, J. J., Robertson, W. J., Spring, S., Schumann, P., Franzmann, P. D. & Puhakka, J. A. (2006).** Novel thermophilic sulfate-reducing bacteria from a geothermally active underground mine in Japan. *Appl Environ Microbiol* **72**, 3759–3762.
- Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H., Yi, H., Won, S. & Chun, J. (2012).** Introducing EzTaxon: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* **62**, 716–721.
- Kimura, M. (1980).** A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Kumar, N. R., Nair, S., Langer, S., Busse, H.-J. & Kämpfer, P. (2008).** *Altererythrobacter indicus* sp. nov., isolated from wild rice (*Porteresia coarctata* Tateoka). *Int J Syst Evol Microbiol* **58**, 839–844.
- Kwon, K. K., Woo, J. H., Yang, S.-H., Kang, J. H., Kang, S. G., Kim, S.-J., Sato, T. & Kato, C. (2007).** *Altererythrobacter epoxidivorans* gen. nov., sp. nov., an epoxide hydrolase-active, mesophilic marine bacterium isolated from cold-seep sediment, and reclassification of *Erythrobacter luteolus* Yoon *et al.* 2005 as *Altererythrobacter luteolus* comb. nov. *Int J Syst Evol Microbiol* **57**, 2207–2211.
- Lai, Q., Yuan, J. & Shao, Z. (2009).** *Altererythrobacter marinus* sp. nov., isolated from deep seawater. *Int J Syst Evol Microbiol* **59**, 2973–2976.
- Lane, D. J. (1991).** 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds) *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons, New York, pp 115–175.

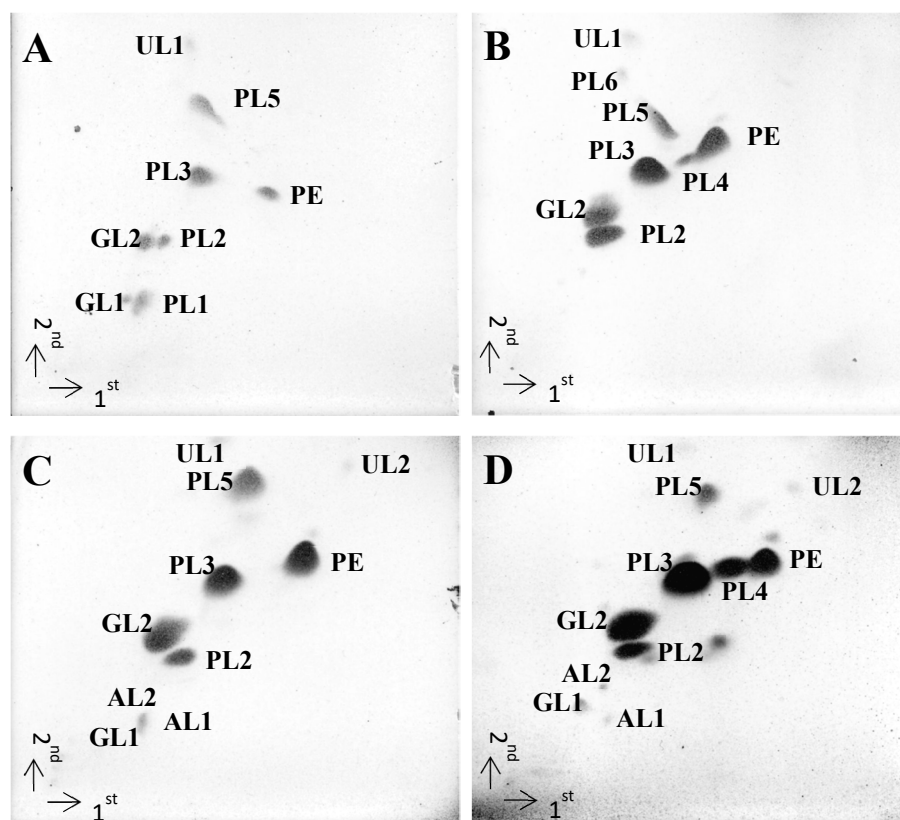


- Lee, K.-B., Liu, C.-T., Anzai, Y., Kim, H., Aono, T. & Oyaizu H. (2005). The hierarchical system of the 'Alphaproteobacteria': description of *Hyphomonadaceae* fam. nov., *Xanthobacteraceae* fam. nov. and *Erythrobacteraceae* fam. nov. *Int J Syst Evol Microbiol* **55**, 1907–1919.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **41**, (Web Server issue):W597-600 doi:10.1093/nar/gkt376.
- Mesbah, M., Premachandran, U. & Whitman, W.B. (1989). Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* **39**, 159–167.
- Park, S., Jung, Y.-T., Park, J.-M. & Yoon, J. H. (2016). *Altererythrobacter confluentis* sp. nov., isolated from water of an estuary environment. *Int J Syst Evol Microbiol* **66**, 4002–4008.
- Park, S. C., Baik, K. S., Choe, H. N., Lim, C. H., Kim, H. J., Ka, J.-O. & Seong, C. N. (2011). *Altererythrobacter namhicola* sp. nov. and *Altererythrobacter aestuarii* sp. nov., isolated from seawater. *Int J Syst Evol Microbiol* **61**, 709–715.
- Proença, D. N., Nobre, M. F. & Morais, P. V. (2014). *Chitinophaga costaii* sp. nov., an endophyte of *Pinus pinaster*, and emended description of *Chitinophaga niabensis*. *Int J Syst Evol Microbiol* **64**, 1237–1243.
- Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Seo, S. H. & Lee, S. D. (2010). *Altererythrobacter marensis* sp. nov., isolated from seawater. *Int J Syst Evol Microbiol* **60**, 307–311.
- Stackebrandt, E. & Goebel, B. M. (1994). Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* **44**, 846–849.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. & Trüper, H. G. (1987). Report of the Ad Hoc Committee on reconciliation of approaches to bacterial systematics. *Int J Syst Evol Microbiol* **37**, 463–464.
- Xue, H., Piao, C.-G., Guo, M.-W., Wang, L.-F., Fang, W. & Li, Y. (2016). Description of *Altererythrobacter aerius* sp. nov., isolated from air, and emended description of the genus *Altererythrobacter*. *Int J Syst Evol Microbiol* **66**, 4543–4548.
- Xue, X., Zhang, K., Cai, F., Dai, J., Wang, Y., Rahman, E., Peng, F. & Fang, C. (2012). *Altererythrobacter xinjiangensis* sp. nov., isolated from desert sand, and emended description of the genus *Altererythrobacter*. *Int J Syst Evol Microbiol* **62**, 28–32.

## Supplementary material



**Supplementary Figure S4.4** Neighbour-joining (NJ) phylogenetic tree based on nearly-full length 16S rRNA gene sequences of strains CPA5<sup>T</sup> and BR75<sup>T</sup>, isolated from *Halimione portulacoides*, and type strains of the genus *Altererythrobacter*, other genera of the family *Erythrobacteraceae*, and relevant species of the *Novosphingobium* genus. Full circles denote nodes that were also recovered in the maximum-likelihood (ML) tree, based on the same sequences. Bootstrap values (expressed as percentages of 1000 replications) which are  $\geq 50\%$  are shown at branch points, following the order: NJ/ML. Accession numbers for the type strains are shown in parenthesis. Bar, 0.01 nt substitution rate ( $K_{nuc}$ ) units.



**Supplementary Figure S4.5** Two-dimensional TLC of total polar lipids of strain CPA5<sup>T</sup> (A), BR75<sup>T</sup> (B), *A. aestuarii* KCTC 22735<sup>T</sup> (C) and *A. marensis* KCTC 22370<sup>T</sup> (D) stained with 5% ethanolic molybdophosphoric acid. PE, phosphatidylethanolamine; AL1-2, unidentified aminolipids; GL1-2, unidentified glycolipids; PL1-6, unidentified phospholipids; UL1-2, unidentified unknown lipids.

**4.4 *Zunongwangia endophytica* sp. nov., an endophyte isolated from the salt marsh plant *Halimione portulacoides*, and emended description of the genus *Zunongwangia***

**Authors**

Cátia Fidalgo<sup>1</sup>, Ricardo Martins<sup>1</sup>, Diogo Neves Proença<sup>2</sup>, Paula V. Morais<sup>2,3</sup>, Artur Alves<sup>1</sup>, Isabel Henriques<sup>1</sup>

<sup>1</sup> Departamento de Biologia, CESAM, Universidade de Aveiro, Aveiro, Portugal

<sup>2</sup> CEMUC, University of Coimbra, 3030-788 Coimbra, Portugal

<sup>3</sup> Department of Life Sciences, FCTUC, University of Coimbra, 3000-456 Coimbra, Portugal

**Publication status**

In press

Int J Syst Evol Microbiol

doi: 10.1099/ijsem.0.002069

## Abstract

Taxonomical analyses were performed on strain CPA58<sup>T</sup>, an isolate obtained from surface-sterilized aboveground tissues of the halophyte *Halimione portulacoides*, collected from a salt marsh in Ria de Aveiro, Portugal. The strain was Gram-stain-negative, rod-shaped, oxidase-negative and catalase-positive. Optimal growth was observed at 26 °C, at pH 6–8 and in the presence of 2 to 3% (w/v) NaCl. Phylogenetic analyses based on the 16S rRNA gene sequence showed that strain CPA58<sup>T</sup> belonged to the genus *Zunongwangia*, with highest sequence similarity to both *Zunongwangia profunda* SM-A87<sup>T</sup> and *Zunongwangia mangrovi* P2E16<sup>T</sup> (96.5 %), followed by *Zunongwangia atlantica* 22II14-10F7<sup>T</sup> (95.9 %). The principal fatty acids were iso-C<sub>15:0</sub>, summed feature 3 (C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH) and iso-C<sub>17:0</sub> 3-OH. The major respiratory quinone was MK-6 and the DNA G+C content was 35.1 mol%. Phylogenetic and chemotaxonomic analyses clearly placed strain CPA58<sup>T</sup> in the genus *Zunongwangia*. Nevertheless, 16S rRNA gene sequence analysis evidenced that the threshold for same species relatedness was not surpassed, and biochemical tests revealed diagnostic characteristics that differentiate this strain from other type strains in the genus. Overall, the analyses performed showed that strain CPA58<sup>T</sup> represents a novel species within the genus *Zunongwangia*, for which the name *Zunongwangia endophytica* sp. nov. is proposed, with the type strain CPA58<sup>T</sup> (= CECT 9128<sup>T</sup> = LMG 29517<sup>T</sup>).

## Keywords

*Flavobacteriaceae*, endophytic, salt marsh, halophytes, taxonomy

## Main text

The genus *Zunongwangia* was proposed by Qin et al. (2007) and belongs to the family *Flavobacteriaceae*. At the time of writing, the genus included three species, isolated from different environments: *Zunongwangia profunda* from deep-sea sediment (Qin et al., 2007), *Zunongwangia atlantica* from rhizosphere of a mangrove tree (Shao et al., 2014) and *Zunongwangia mangrovi* from deep-sea water (Rameshkumar et al., 2014). The genus is described as comprising Gram-negative, non-motile and non-spore forming, strictly aerobic rods, catalase- and oxidase-positive, without flexirubin, with MK-6 as the main menaquinone, and with the following main fatty acids: iso-C<sub>15:0</sub>, summed feature 3 (iso-C<sub>15:0</sub> 2-OH and/or C<sub>16:1</sub> ω7c), iso-C<sub>17:0</sub> 3-OH, iso-C<sub>15:1</sub> G, iso-C<sub>17:1</sub> ω9c, C<sub>15:0</sub>, iso-C<sub>15:0</sub> 3-OH (Qin et al., 2007). An emendation to the description of the genus added that some species require sodium ions for growth (Rameshkumar et al., 2014).

In the scope of a diversity study of endophytic bacteria of the halophyte *Halimione portulacoides* from a salt marsh in Ria de Aveiro, Portugal, 665 isolates were obtained and characterized (Fidalgo et al., 2016). The present study focuses on one of these isolates, designated strain CPA58<sup>T</sup>, isolated from aboveground tissues of *H. portulacoides*. Comparative 16S rRNA gene sequence analyses showed that strain CPA58<sup>T</sup> was closely related to the genus *Zunongwangia*. The aim of this study was to determine the exact taxonomic position of strain CPA58<sup>T</sup> using a polyphasic approach, as suggested by the minimal standards for describing new taxa in the family *Flavobacteriaceae* (Bernardet et al., 2002).

Healthy specimens of the halophyte *Halimione portulacoides* were collected from a salt marsh in September of 2012. Aboveground and belowground tissues from these specimens were surface-sterilized, macerated in phosphate buffer solution and studied for their bacterial diversity. The bacterial isolation was performed on three culture media: Tryptic Soy Agar (TSA, Merck, Germany), Marine Agar (MA, Difco Laboratories, France) and R2A (Merck, Germany) (Fidalgo et al., 2016). Strain CPA58<sup>T</sup> was originally isolated from and routinely cultured on MA culture medium, at 28 °C, under aerobic conditions.

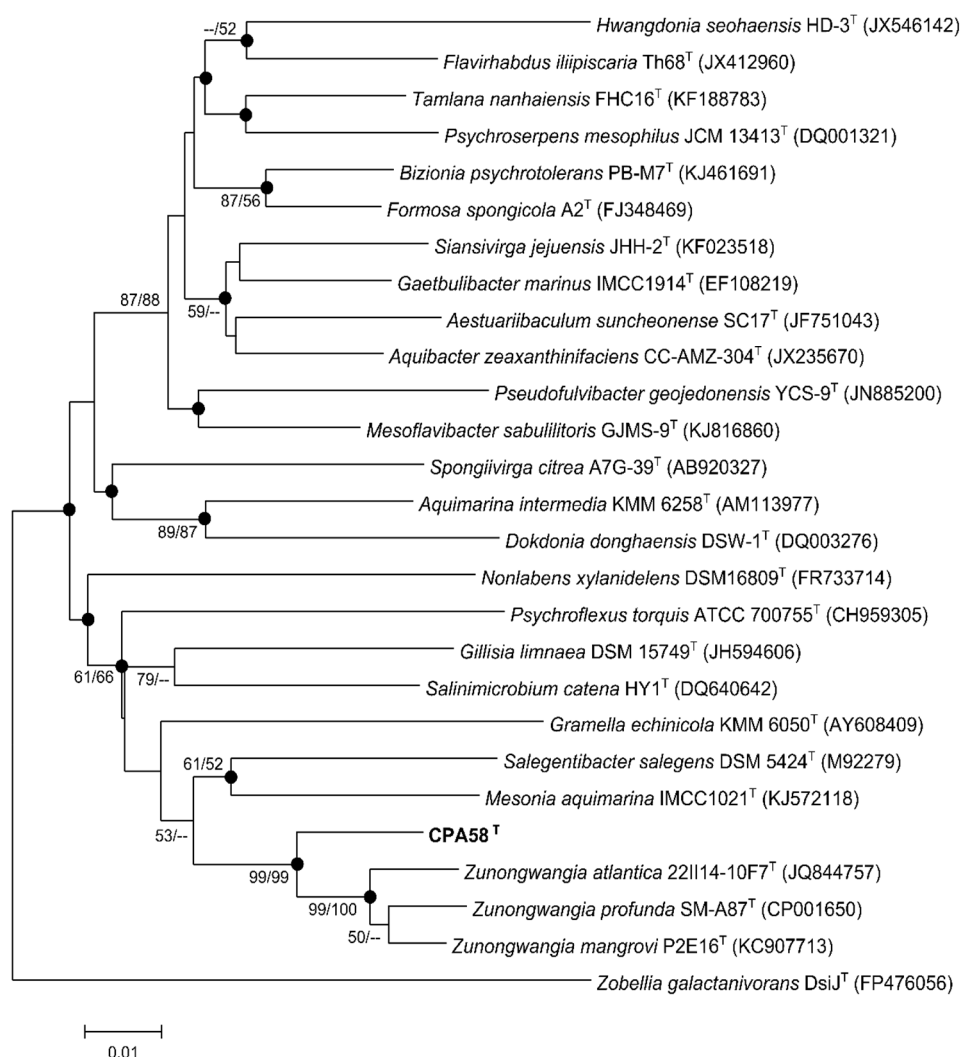
Genomic DNA was prepared using the Genomic DNA Purification kit #0513 (Thermo Scientific, USA) according to the manufacturer's instructions. Amplification of the 16S rRNA gene was performed as described elsewhere (Fidalgo et al., 2016) and sequencing was done with the universal primers 27F (Lane, 1991) and 704F (Kaksonen et al., 2006). Sequences of related taxa were obtained from

the EzTaxon database (Yoon et al., 2016). Sequence alignment was performed with Clustal Omega (McWilliam et al., 2013), and edited with BioEdit version 7.2.5 (Hall, 1999). Phylogenetic analyses were performed using MEGA version 6.0 (Tamura et al., 2013). The Kimura two-parameter model (Kimura, 1980) was used in clustering with the neighbour-joining (NJ, Saitou & Nei, 1987) and maximum-likelihood (ML, Felsenstein, 1981). Bootstrap values based on 1000 replications were obtained in these phylogenetic analyses.

The near full-length 16S rRNA gene sequence of strain CPA58<sup>T</sup> was obtained (1429 nt). The 16S rRNA gene sequence similarity was highest (96.5 %) with *Zunongwangia profunda* SM-A87<sup>T</sup> and *Zunongwangia mangrovi* P2E16<sup>T</sup>, followed by 95.9 % with *Zunongwangia atlantica* 22II14-10F7<sup>T</sup> and 95.03 % with *Salegentibacter chungangensis* CAU 1289<sup>T</sup>, while the similarities to other type strains were all below 94.0 %. These results support that strain CPA58<sup>T</sup> belongs to the genus *Zunongwangia*. A phylogenetic tree was built, based on 16S rRNA gene sequences of strain CPA58<sup>T</sup> and related type strains. As presented in Figure 4.6, CPA58<sup>T</sup> forms a clear monophyletic cluster with other *Zunongwangia* type strains, with high bootstrap support (99 %).

Optimal growth conditions were assessed in MA medium. Ability to grow at different temperatures was tested by incubating at 4, 18, 26, 30, 37, 42.5 and 50 °C. At optimal growth temperature, the pH range for growth was tested from pH 4 to 12 in 1 unit intervals. The following buffers were used to adjust the pH values: MES (4-Morpholineethanesulfonic acid, pH 4 to 6; Sigma, USA), Tricine (pH 7 and 8; Sigma, USA), and CAPS (N-cyclohexyl-3-aminopropanesulfonic acid, pH 9 to 12; Sigma, USA). At optimal growth temperature and pH, tolerance to NaCl was tested using MA at ¼ strength as the basal medium, to test NaCl concentrations of 0.5, 1 [adding 0.5 % NaCl (w/v)], 2, 3, 5, 10, 15 and 20 % (w/v).

Unless otherwise stated, cells were grown on MA supplemented with 1 % (w/v) NaCl, incubated at 26 °C for 2 days. Gram staining was performed using a Gram staining set (Merck, Germany) according to the manufacturer's instructions. Cell morphology and motility were observed using light microscopy (Nikon 80i). Gliding motility was assessed after growing the strain in half strength MA for 72h, using cavity slides and following the hang-drop method described in Bowman (2000). The presence of flexirubin-type pigments was assessed as described by Bowman (2000). Catalase activity was determined using the H<sub>2</sub>O<sub>2</sub>-reagent (Liofilchem, Italy). Oxidase activity was evaluated using oxidase strips (Liofilchem, Italy).



**Figure 4.5** Neighbour-joining (NJ) tree showing the phylogenetic positions of strain CPA58<sup>T</sup> and representatives of other related taxa based on 16S rRNA gene sequences. Filled circles indicate nodes that were also recovered in maximum-likelihood (ML) tree, based on the same sequences. Bootstrap values (expressed as percentages of 1000 replications)  $\geq 50\%$  are shown at branch points for NJ/ML trees. *Zobellia galactanivorans* Dsij<sup>T</sup> was used as an outgroup. Accession numbers for the type strains are shown in parenthesis. Bar, 0.01 nt substitution rate ( $K_{nuc}$ ) units.

Anaerobic growth was assessed using thioglycollate medium (Merck, Germany) supplemented with 2 % NaCl, and growth was monitored for 7 days. DNA hydrolysis was determined on DNase agar (HiMedia, India) with 0.1 % toluidine blue, and supplemented with 2 % NaCl. Citrate utilization was assessed using Simmon's Citrate Agar (Oxoid, USA) supplemented with 2 % NaCl. Production of H<sub>2</sub>S was evaluated using Kligler's Iron Agar (Merck, Germany) supplemented with 2 % NaCl. Hydrolysis



of casein, starch, tween 20, cellulose and xylan, phosphate solubilization, siderophore production, 1-aminocyclopropane-1-carboxylate (ACC) deaminase activity and indole acetic acid (IAA) production were assessed as described by Fidalgo et al. (2016).

Strain CPA58<sup>T</sup> was Gram-stain-negative, oxidase-negative, catalase-positive, non-pigmented, rod-shaped and non-motile. Growth in thioglycollate medium showed a microaerophilic behavior. Other physiological characteristics of strain CPA58<sup>T</sup> are stated in the species description section.

Biochemical tests were carried out using API 20NE, API ZYM and API 50CH strips (bioMérieux, France) according to the manufacturer's instructions, except for using 0.9% (w/v) NaCl solution for preparing cell suspension. For strain CPA58<sup>T</sup> cell suspension was prepared using a 4 % (w/v) sea salts (Sigma, USA) solution in order to accommodate the salinity requirements of this strain. *Zunongwangia profunda* DSM 18752<sup>T</sup> and *Z. mangrovi* DSM 24499<sup>T</sup> were tested for comparison. These results are given in the species description and Table 4.9.

Determination of G+C content was performed by HPLC (Mesbah et al., 1989). The DNA G+C content of the endophytic isolate CPA58<sup>T</sup> was 35.1 mol%. The value is similar to that of *Z. profunda* SM-A87<sup>T</sup> (36.2 %) and *Z. mangrovi* P2E16<sup>T</sup> (34.3 %) that were obtained from genome data (Qin et al., 2010; Rameshkumar et al., 2014).

For lipoquinone analysis, strain CPA58<sup>T</sup> was grown on marine broth (MB) at 26 °C for 72 h, and the cells were then harvested and lyophilized. Lipoquinones were then extracted and analyzed as described by Proença et al. (2014). The major respiratory quinone of strain CPA58<sup>T</sup> was determined to be MK-6, which is in accordance with the characteristics of the genus *Zunongwangia*.

For fatty acid analysis, cells were grown on MA at 26 °C for 48h, and the analysis was conducted as described by Proença et al. (2014). The Agilent Technologies 6890N Network GC System and Standard MIS Library Generation Software (SherlockMicrobial ID System, TSBA 6 database, version 6.0; USA) were used. The fatty acids profile of CPA58<sup>T</sup> was produced in parallel with those of *Z. profunda* DSM 18752<sup>T</sup> and *Z. mangrovi* DSM 24499<sup>T</sup>. The major fatty acids of strain CPA58<sup>T</sup> were iso-C<sub>15:0</sub>, summed feature 3 (iso-C<sub>15:0</sub> 2-OH and/or C<sub>16:1</sub> ω7c) and iso-C<sub>17:0</sub> 3-OH, which accounted for over 45 % of the total fatty acids. The results are presented in Table 4.10 and are in accordance with what is described for the genus *Zunongwangia*.

**Table 4.9** Physiological characteristics of strain CPA58<sup>T</sup> and recognized species of the genus *Zunongwangia* (data from this study unless otherwise stated).

Characteristic	CPA58 <sup>T</sup>	<i>Z. mangrovi</i> DSM 24499 <sup>T</sup>	<i>Z. profunda</i> DSM 18752 <sup>T</sup>
Oxidase	–	+ <sup>a</sup>	+ <sup>b</sup>
Relation to oxygen	Microaerophilic	Microaerophilic	Microaerophilic
Hydrolysis of:			
Casein	–	+ <sup>a</sup>	+ <sup>b</sup>
Starch	–	+ <sup>a</sup>	– <sup>b</sup>
Temperature range	4–30 °C	20–37 °C <sup>a</sup>	4–38 °C <sup>b</sup>
pH range	6–9	6–9 <sup>a</sup>	5–8.5 <sup>b</sup>
NaCl (w/v) range	2–15 %	0.3–12 % <sup>a</sup>	0–12 % <sup>b</sup>
API 20NE results:			
Reduction of nitrates to nitrites	–	–	– †
Hydrolysis of gelatin	+	–	–
<i>N</i> -Acetylglucosamine assimilation	–	+	–
Potassium gluconate assimilation	–	w	–
API ZYM results:			
Esterase (C4), esterase lipase (C8)	+	+	+ †
Valine arylamidase, $\beta$ -glucosidase	+	+ *	+
Cystine arylamidase, $\alpha$ -chymotrypsin, $\alpha$ -galactosidase	w	+ *	+ †
$\beta$ -Galactosidase	w	+ *	+
Lipase (C14)	–	w	w
$\beta$ -Glucuronidase	–	w *	–
$\alpha$ -Mannosidase	–	–	w †
$\alpha$ -Fucosidase	–	+ *	–
API 50CH results:			
Glycerol, D-fructose, L-rhamnose, D-trehalose	–	– *	– †
L-Arabinose, D-galactose, D-mannose, D-sorbitol, D-cellobiose, D-maltose, D-melibiose, D-saccharose (sucrose), D-raffinose	–	–	– †
D-Mannitol	–	–	+
DNA G+C content (mol%)	35.1	34.3 <sup>a</sup>	35.8 <sup>b</sup>

+, Positive; w, weakly positive; –, negative; nd, not determined; <sup>a</sup>, data from Rameshkumar et al. (2014); <sup>b</sup>, data from Qin et al. (2007); \*, results differed from those published in Rameshkumar et al. (2014); †, results differed from those published in Qin et al. (2007).

**Table 4.10** Fatty acid composition of strains CPA58<sup>T</sup>, *Z. profunda* DSM 18752<sup>T</sup>, *Z. mangrovi* DSM 24499<sup>T</sup>.

Fatty acid	CPA58 <sup>T</sup>	<i>Z. profunda</i> DSM 18752 <sup>T</sup>	<i>Z. mangrovi</i> DSM 24499 <sup>T</sup>
Unknown ECL			
13.563	7.6	6.5	6.7
13.652	1.0	tr	–
15.938	1.2	1.9	tr
Saturated			
C <sub>15:0</sub>	6.6	4.4	6.9
Unsaturated			
C <sub>15:1</sub> ω6c	1.4	tr	tr
C <sub>17:1</sub> ω6c	1.9	1.2	3.0
Branched-chain			
iso-C <sub>15:0</sub>	20.8	16.3	22.4
iso-C <sub>15:0</sub> 3-OH	2.8	3.4	2.9
iso-C <sub>15:1</sub> G	7.4	10.4	9.8
iso-C <sub>16:0</sub>	1.3	1.4	3.1
iso-C <sub>16:0</sub> 3-OH	1.4	1.3	1.1
iso-C <sub>17:0</sub>	tr	tr	1.2
iso-C <sub>17:0</sub> 3-OH	11.7	12.4	11.0
anteiso-C <sub>15:0</sub>	8.4	5.3	3.9
anteiso-C <sub>17:1</sub> ω9c	1.4	–	–
Hydroxyl			
C <sub>15:0</sub> 2-OH	2.0	1.9	1.3
C <sub>15:0</sub> 3-OH 12-methyl/C <sub>16:0</sub>	–	1.4	1.3
C <sub>17:0</sub> 2-OH	1.6	tr	tr
Summed feature			
3	13.5	21.3	12.3
9	3.8	3.6	5.3

ECL, equivalent chain length; –, not detected; tr, trace (< 1 %). Summed feature 3 contains C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH; Summed feature 9 contains iso-C<sub>17:1</sub> ω9c and/or 10-methyl-C<sub>16:0</sub>.

The differences in physiological, biochemical and chemotaxonomic characteristics between strain CPA58<sup>T</sup>, *Z. profunda* DSM 18752<sup>T</sup> and *Z. mangrovi* DSM 24499<sup>T</sup> are given in Table 4.9. The 16S rRNA gene sequence similarity between strain CPA58<sup>T</sup> and *Z. profunda* SM-A87<sup>T</sup> and *Z. mangrovi* P2E16<sup>T</sup> strongly indicates that strain CPA58<sup>T</sup> belongs to the genus *Zunongwangia*. Strain CPA58<sup>T</sup> represents, however, a novel and distinct species since it can be differentiated from other type strains of the genus on the basis of physiological and chemotaxonomic characteristics. Additionally, organisms that present a 16S rRNA gene sequence similarity below 97 % rarely display a DNA-DNA homology greater than 70 % (Stackebrandt & Goebel, 1994), which is below the threshold for same species

relatedness (Wayne et al., 1987). On the basis of the data described above, strain CPA58<sup>T</sup> should be regarded as representing a novel species of the genus *Zunongwangia*, for which the name *Zunongwangia endophytica* sp. nov. is proposed. An emended description of the genus *Zunongwangia* is also proposed.

#### **Emended description of the genus *Zunongwangia* Qin et al. 2007, emend. Rameshkumar et al. 2014**

The description of the genus *Zunongwangia* is as given by Qin et al. (2007) and Rameshkumar et al. (2014) with the following amendment: species are microaerophilic or strictly aerobic, and variable for oxidase reaction.

#### **Description of *Zunongwangia endophytica* sp. nov.**

*Zunongwangia endophytica* (en.do.phy'ti.ca. Gr. pref. *endo* within; Gr. n. *phyton* plant; L. fem. suff. -ica adjectival suffix used with the sense of belonging to; N.L. fem. adj. *endophytica* within plant, pertaining to the endophytic nature of the strain and its isolation from internal plant tissues).

Cells are Gram-stain-negative, rod-shaped, 0.55–1.13 µm wide and 1.38–3.4 µm long and non-motile. Positive for catalase, phosphate solubilization, production of IAA, xylan hydrolysis and hydrolysis of Tween 20, but negative for oxidase, siderophore production, ACC deaminase activity, DNA hydrolysis, citrate utilization, H<sub>2</sub>S production, and hydrolysis of starch, carboxymethyl cellulose and casein (hydrolysis of skim milk). Cannot grow under anaerobic conditions, but presents microaerobic behavior. Flexirubin-type pigments were absent. On MA plates forms smooth and creamy yellow colonies with smooth regular edges, of 0.6–2.2 mm in diameter after incubating at 26 °C for 48 h. Slightly halophilic, grows in the presence of 2–10% NaCl (optimum 2–3 %), at 4–30 °C (optimum 26 °C) and at pH 6–9 (optimum 6–8). The main fatty acids are iso-C<sub>15:0</sub>, iso-C<sub>17:0</sub> 3-OH, and summed feature 3 (iso-C<sub>15:0</sub> 2-OH and/or C<sub>16:1</sub> ω7c). The major respiratory quinone is MK-6. In the API 20NE test, it can hydrolyze esculin (β-glucosidase), gelatin (protease) and para-nitrophenyl-β-D-galactopyranose (β-galactosidase) but is negative for fermentation of D-glucose, activity of arginine dihydrolase, urease, indole production, reduction of nitrates to nitrites and to nitrogen, assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetylglucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid. In the API ZYM test strip, results are positive for alkaline phosphatase, esterase (C4), esterase lipase

(C8), leucine arylamidase, valine arylamidase, trypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase,  $\alpha$ -glucosidase,  $\beta$ -glucosidase, *N*-acetyl- $\beta$ -glucosaminidase; weakly positive for cystine arylamidase,  $\alpha$ -chymotrypsin,  $\alpha$ -galactosidase and  $\beta$ -galactosidase; negative for lipase (C14),  $\beta$ -glucuronidase,  $\alpha$ -mannosidase and  $\alpha$ -fucosidase. In the API 50CH test strip, CPA58<sup>T</sup> tested positive for acid production from aesculin ferric citrate, and negative for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl- $\beta$ -D-xylopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl- $\alpha$ -D-mannopyranoside, methyl- $\alpha$ -D-glucopyranoside, *N*-acetylglucosamine, amygdalin, arbutin, salicin, D-celiobiose, D-maltose, D-lactose (bovine), D-melibiose, D-saccharose (sucrose), D-trehalose, inulin, D-melezitose, D-raffinose, amidon (starch), glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate, and potassium 5-ketogluconate.

The type strain, CPA58<sup>T</sup> (= CECT 9128<sup>T</sup> = LMG 29517<sup>T</sup>) was isolated from the surface-sterilized tissues of the halophyte *Halimione portulacoides*. The G+C content of the DNA of the type strain is 35.1 mol%.

## Acknowledgements

This work was supported by European Funds (FEDER) through COMPETE and by National Funds through the Portuguese Foundation for Science and Technology (FCT) within project PhytoMarsh (PTDC/AAC-AMB/118873/2010 – FCOMP-01-0124-FEDER-019328) and research unit CESAM (UID/AMB/50017/2013 – POCI-01-0145-FEDER-007638). FCT is acknowledged for financing to Artur Alves (FCT Investigator Programme – IF/00835/2013), Isabel Henriques (FCT Investigator Programme – IF/00492/2013), Diogo Proença (post doc grant – SFRH/BPD/100721/2014) and Cátia Fidalgo (PhD grant – SFRH/BD/85423/2012).

## References

- Bernardet, J. F., Nakagawa, Y., Holmes, B. & Subcommittee on the taxonomy of *Flavobacterium* and *Cytophaga*-like bacteria of the International Committee on Systematics of Prokaryotes (2002). Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. *Int J Syst Evol Microbiol* **52**, 1049–1070.
- Bowman, J. P. (2000). Description of *Cellulophaga algicola* sp. nov., isolated from the surfaces of Antarctic algae, and reclassification of *Cytophaga uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Cellulophaga uliginosa* comb. nov. *Int J Syst Evol Microbiol* **50**, 1861–1868.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum-likelihood approach. *J Mol Evol* **17**, 368–376.
- Fidalgo, C., Henriques, I., Rocha, J., Tacão, M. & Alves, A. (2016). Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization, and influence of metal(loid) contamination. *Environ Sci Pollut Res* **23**(10), 10200-10214.
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment edit and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* **41**, 95–98.
- Kaksonen, A. H., Plumb, J. J., Robertson, W. J., Spring, S., Schumann, P., Franzmann, P. D. & Puhakka, J. A. (2006). Novel thermophilic sulfate-reducing bacteria from a geothermally active underground mine in Japan. *Appl Environ Microbiol* **72**(5), 3759-3762.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Lane, D. J. (1991). 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*, pp. 115–175. Edited by E. Stackebrandt & M. Goodfellow. New York, NY: John Wiley and Sons.
- Mesbah, M., Premachandran, U. & Whitman, W.B. (1989). Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* **39**, 159–167.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **41**, (Web Server issue):W597-600 doi:10.1093/nar/gkt376.
- Proença, D. N., Nobre, M. F. & Morais, P. V. (2014). *Chitinophaga costaii* sp. nov., an endophyte of *Pinus pinaster*, and emended description of *Chitinophaga niabensis*. *Int J Syst Evol Microbiol* **64**, 1237-1243.

- Qin, Q. L., Zhao, D. L., Wang, J., Chen, X. L., Dang, H. Y., Li, T. G., Zhang, Y. Z. & Gao, P. J. (2007).** *Wangia profunda* gen. nov., sp. nov., a novel marine bacterium of the family *Flavobacteriaceae* isolated from southern Okinawa Trough deep-sea sediment. *FEMS Microbiol Lett* **271**, 53–58.
- Qin, Q. L., Zhang, X. Y., Wang, X. M., Liu, G. M., Chen, X. L., Xie, B. B., Dang, H. Y., Zhou, B. C., Yu, J. & Zhang, Y. Z. (2010).** The complete genome of *Zunongwangia profunda* SM-A87 reveals its adaptation to the deep-sea environment and ecological role in sedimentary organic nitrogen degradation. *BMC Genomics* **11**, 247.
- Rameshkumar, N., Krishnan, R., Lang, E., Matsumura, Y., Sawabe, T. & Sawabe, T. (2014).** *Zunongwangia mangrovi* sp. nov., isolated from mangrove (*Avicennia marina*) rhizosphere, and emended description of the genus *Zunongwangia*. *Int J Syst Evol Microbiol* **64**, 545–550.
- Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Shao, R., Lai, Q., Liu, X., Sun, F., Du, Y., Li, G. & Shao, Z. (2014).** *Zunongangia atlantica* sp. nov., isolated from deep-sea water. *Int J Syst Evol Microbiol* **64**, 16–20.
- Stackebrandt, E. & Goebel, B. M. (1994).** Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* **44**, 846–849.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013).** MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. & Trüper, H. G. (1987).** Report of the Ad Hoc Committee on reconciliation of approaches to bacterial systematics. *Int J Syst Evol Microbiol* **37**, 463–464.
- Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J. (2016).** Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. *Int J Syst Evol Microbiol*; doi: 10.1099/ijsem.0.001755

**4.5 The endosphere of the salt marsh plant *Halimione portulacoides* is a diversity hotspot for the genus *Salinicola*: description of five novel species *Salinicola halimionae* sp. nov., *Salinicola aestuarina* sp. nov., *Salinicola endophytica* sp. nov., *Salinicola halophytica* sp. nov. and *Salinicola lusitana* sp. nov.**

## **Authors**

Cátia Fidalgo<sup>1</sup>, Diogo Neves Proença<sup>2</sup>, Paula V. Morais<sup>2,3</sup>, Isabel Henriques<sup>1</sup>, Artur Alves<sup>1</sup>

<sup>1</sup> Departamento de Biologia, CESAM, Universidade de Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

<sup>2</sup> CEMMPRE, University of Coimbra, 3030-788 Coimbra, Portugal

<sup>3</sup> Department of Life Sciences, FCTUC, University of Coimbra, 3000-456 Coimbra, Portugal

## **Publication status**

Submitted



## Abstract

Seven motile Gram-stain negative endophytic strains were isolated from surface-sterilized tissues of the halophyte *Halimione portulacoides* collected from a salt marsh in Ria de Aveiro, Portugal. These strains were analyzed in order to determine their exact taxonomic position, using five closely related type strains of *Salinicola* spp. in comparative analyses. Genome sequencing and its comparison indicated that five of the seven isolated strains comprised five distinct and novel species (average nucleotide identity < 0.95; *in silico* DNA-DNA hybridization < 70 %; G+C difference > 1 %). A multilocus sequence analysis was performed using *gyrB*, *rpoD* and 16S rRNA gene sequences from endophytic strains from the present study and type strains to determine their phylogenetic positions. The novel strains are facultative anaerobes, mesophilic, facultative alkaliphilic and halophilic, test positive for catalase and oxidase activities, for hydrolysis of Tween 20 and phosphate, production of IAA, but do not produce H<sub>2</sub>S. Ubiquinone Q-9 is present in major amounts in all strains and the major fatty acids included C<sub>16:0</sub> and the summed feature containing C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c. The DNA G+C content ranged from 60.6 to 65.8 mol%. Genomic, biochemical and physiological characteristics confirmed five strains as new species belonging to the genus *Salinicola*, for which the names *Salinicola halimionae* sp. nov. (type strain CPA60<sup>T</sup> = CECT 9338<sup>T</sup> = LMG 30107<sup>T</sup>), *Salinicola aestuarina* sp. nov. (type strain CPA62<sup>T</sup> = CECT 9339<sup>T</sup> = LMG 30108<sup>T</sup>), *Salinicola endophytica* sp. nov. (type strain CPA92<sup>T</sup> = CECT 9340<sup>T</sup> = LMG 30109<sup>T</sup>), *Salinicola halophytica* sp. nov. (type strain CR45<sup>T</sup> = CECT 9341<sup>T</sup> = LMG 30105<sup>T</sup>) and *Salinicola lusitana* sp. nov. (type strain CR50<sup>T</sup> = CECT 9342<sup>T</sup> = LMG 30106<sup>T</sup>) are proposed.

## Keywords

*Halomonadaceae*, endophytic, salt-marsh, halophytes, taxonomy

## Background

### *Halomonadaceae* family

The family *Halomonadaceae* belongs to the class Gammaproteobacteria and was proposed by Franzmann et al. (1988). This family was initially described as harboring non-spore forming Gram-negative oxidase-positive rods that require NaCl for growth. Ubiquinones were determined to be the respiratory lipoquinone class present in this family, and G+C content ranged from 52 to 68 mol%. At the time, the family was described as including the genera: *Halomonas*, the type genus of the family, and *Deleya* (Franzmann et al., 1988). Subsequent analyses of the 16S rRNA gene sequence lead to addition of the genera *Chromohalobacter* and *Volcaniella* in the *Halomonas-Deleya* cluster, with reclassification of the type species of the genus *Volcaniella* as pertaining to the genus *Halomonas* (Mellado et al., 1995). The following year, species of the genera *Deleya* and *Halovibrio*, as well as the species *Paracoccus halodenitrificans*, were reclassified as members of the genus *Halomonas* (Dobson & Franzmann, 1996). These authors also proposed the inclusion of the monospecific genus *Zymobacter* (Okamoto et al., 1993) in the family *Halomonadaceae*.

The fast growing genus *Halomonas* soon demanded a more extensive analysis to understand the relationships of its different species. Mata et al. (2002) performed a detailed phenotypic characterization of the 21 *Halomonas* species described at the time, detailing and citing many methodologies that are still currently used in the description of new taxa of the family *Halomonadaceae*. Further phylogenetic analyses based on 16S and 23S rRNA gene sequences of the three genera of the family allowed clarification of early confusion on phylogenetic positioning of species in the family *Halomonadaceae*: the phylogenetic positions of species of the genera *Chromohalobacter* and *Zymobacter* were confirmed, and it was determined that the genus *Halomonas* comprehended two distinguishable phylogenetic groups (Arahal et al., 2002a). The family *Halomonadaceae* grew as novel genera were described: *Carnimonas* (Arahal et al., 2002a), *Cobetia* (Arahal et al., 2002b), *Halotalea* (Ntougias et al., 2007), *Modicisalibacter* (Gam et al., 2007), *Salinicola* (Anan'ina et al., 2007), *Aidingimonas* (Wang et al., 2009) and *Larsenia* (León et al., 2014). Reclassifications continued to occur as the genus *Cobetia* was described by reassigning a species previously included in the genus *Halomonas* (Arahal et al., 2002b). In 2009, Sánchez-Porro et al. described the novel genus *Kushneria* and proposed the reclassification of three *Halomonas* species into this new genus. More reclassifications occurred in 2010, when the genus *Salinicola* accommodated two additional species, one previously included in the genus *Halomonas* and another previously included in the genus *Chromohalobacter* (de la Haba et al., 2010b).

Early on meetings of the Subcommittee on the Taxonomy of the *Halomonadaceae* family stressed the importance of the determination of recommended minimal standards for the description of new taxa in the family and the importance of using a high quality full or near full length sequence of 16S rRNA gene (Vreeland & Ventosa, 2003). A set of required and recommended minimal standards to describe new *Halomonadaceae* members was detailed by Arahal et al. (2007). These authors emphasized the importance of a high quality (< 0.5 % ambiguity) and near full length (> 1400 nt) 16S rRNA gene sequence, and the importance of treeing methods, seen as different methods may result in differences in topologies, and should be evaluated by statistical methods such as bootstrap analysis.

The description of the family has been emended over time, and earlier emends were mainly concerned with updating signature nucleotides in the 16S rRNA gene sequence that were family-specific (Dobson & Franzmann, 1996; Ntougias et al., 2007; Gam et al., 2007). However, de la Haba et al. (2010a) stated that the use of signature nucleotides was an arbitrary feature, and suggested that using the complete or almost complete sequence of the 16S rRNA gene would be more coherent. These authors investigated the phylogenetic relationships among *Halomonadaceae* members based on, not only 16S rRNA gene sequence, but also on the 23S rRNA gene sequence. It was observed that the resolution using 23S rRNA gene sequence was higher than that observed with 16S rRNA gene sequence, however, still not enough to resolve closely related species (de la Haba et al., 2010a). To overcome this issue, an approach based on multi locus sequence analysis (MLSA) of housekeeping genes was later proposed by de la Haba et al. (2012). These authors concluded that the combination of the genes coding for 16S rRNA, DNA gyrase subunit B (*gyrB*) and RNA polymerase sigma 70 factor (*rpoD*) was adequate for MLSA analysis in the family *Halomonadaceae*. In fact, the 2013 update of recommended minimal standards for description of new *Halomonadaceae* taxa included the MLSA approach with the three aforementioned genes (Oren & Ventosa, 2013).

At the time of writing, the most recent Subcommittee on the Taxonomy of the *Halomonadaceae* family recognized 11 genera and a total of 116 species in this family. This publication also informed that the recommended minimal standards include those detailed by Arahal et al. (2007) supplemented by the MLSA approach of de la Haba et al. (2012) (Oren & Ventosa, 2016).

### The genus *Salinicola*

The genus *Salinicola* was introduced in the family *Halomonadaceae* as a monospecific genus (*S. socius*) which encompasses Gram-negative rods that do not form spores, are motile by means of a single polar flagellum, react positively in the O/F test, and are catalase-positive and oxidase-negative (Anan'ina et al., 2007). This monospecific genus was also described as being moderately halophilic, mesophilic, and as presenting optimal growth from 30 to 37 °C (Anan'ina et al., 2007). As described for the family, the predominant ubiquinone was Q-9 and the major fatty acids detected were C<sub>16:1</sub> ω7c, C<sub>16:0</sub>, C<sub>18:1</sub> ω7c, and cyclic-C<sub>19:0</sub>. The G+C content for the type species was determined as 63 mol%, which is included in the range described for the family *Halomonadaceae* (52-68 mol%, Franzmann et al., 1988).

A polyphasic approach with analyses of the 16S rRNA and 23S rRNA gene sequences, and phenotypic and chemotaxonomic characterization allowed reclassification of two species previously assigned to the genera *Halomonas* (*H. salaria*) and *Chromohalobacter* (*C. salarius*) to the genus *Salinicola* as *S. salarius* and *S. halophilus*, respectively (de la Haba et al., 2010b). These authors included an extensive emend to the genus, of which the following features are of relevance: growth is observed in the range of 15 to 40 °C, respiration on fumarate, nitrate and nitrite are not observed, Tween 20 is hydrolyzed, but casein, DNA, tyrosine and aesculin are not (de la Haba et al., 2010b).

Since then, several species have been described in the genus, namely *Salinicola peritrichatus* (Huo et al., 2013), *S. zeshunii* (Cao et al., 2013), *S. acroporae* (Lepcha et al., 2015) and *S. rhizosphaerae* (Raju et al., 2016). These species were mostly isolated from salt-rich environments such as salt mine (*S. socius*), saline water (*S. salarius*), solar saltern (*S. halophilus*), deep-sea sediment (*S. peritrichatus*), corals (*S. acroporae*) and mangrove (*S. rhizosphaerae*).

## Motivation

Our assessment of culturable endophytic bacteria from *Halimione portulacoides* yielded 665 isolates (Fidalgo et al., 2016), seven of which are the focus of the present section. These isolates were obtained from aboveground (CPA60<sup>T</sup>, CPA62<sup>T</sup> and CPA92<sup>T</sup>) and from belowground tissues (CR45<sup>T</sup>, CR50<sup>T</sup>, CR57 and RZ23) of the halophyte, and were identified as *Salinicola* spp. (Fidalgo et al., 2016). A preliminary assessment of the 16S rRNA gene sequence suggested that these isolates could represent novel species of the genus *Salinicola*. Our aim was then to determine the exact taxonomic position of these strains within the family *Halomonadaceae* and fully characterize the putative new *Salinicola* species using a polyphasic approach. To do so, we followed specifications of the recommended minimal standards in Arahal et al. (2007) and the most recent International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of *Halobacteriaceae* and subcommittee on the taxonomy of *Halomonadaceae* (Oren & Ventosa, 2016).

## Methods

### Strains and culture medium

Endophytic *Salinicola* spp. strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup>, CR50<sup>T</sup>, CR57 and RZ23 were firstly isolated and regularly cultured in Marine Agar medium (MA, Difco Laboratories, France). Closely related type strains of five *Salinicola* species were used as reference strains for comparison in phenotypic and molecular characterization tests: *Salinicola socius* DSM 19940<sup>T</sup>, *Salinicola salarius* DSM 18044<sup>T</sup>, *Salinicola acroporae* LMG 28587<sup>T</sup>, *Salinicola halophilus* CECT 5903<sup>T</sup>, and *Salinicola peritrichatus* JCM 18795<sup>T</sup>. The reference type strains were also maintained on MA.

### DNA extraction, genome sequencing and analysis

To accomplish the minimal standards recommended by Arahal et al. (2007) and de la Haba et al. (2012), a MLSA approach was conducted. For this, DNA of endophytic strains and related type strains of the genus *Salinicola* was extracted using the Wizard Genomic DNA Purification Kit (Promega, USA) according to manufacturer's instructions. DNA quality was attested by running the samples on a 0.8 % agarose gel, and DNA quantification was determined using a Qubit fluorimeter

(Thermo Scientific, USA). Whole genome sequencing was conducted at Stabvida (Portugal) using an Illumina HiSeq 2500 platform. The raw sequences were analyzed using CLC Genomics Workbench version 9.5.3 using a 2 nucleotide cutoff for ambiguity limit, a 0.01 quality limit (error probability) and a minimum number of nucleotides of 30. The quality of the data was checked with the FastQC tool (Andrews, 2010) and the sequences were *de novo* assembled. The genomes were annotated with the Rapid Annotation using System Technology (RAST; Aziz et al., 2008; Overbeek et al., 2014) and the rRNA gene sequences were predicted using the RNAmmer1.2 server (Lagesen et al., 2007). The G+C content and gene sequences for 16S rRNA, GyrB and RpoD for endophytic and related type strains of *Salinicola* species were obtained from genomic sequences. Sequences for these genes from the recently published *Salinicola rhizosphaerae* MSSRFH1<sup>T</sup> and from *Halomonas elongata* DSM 2581<sup>T</sup> were obtained from the EzTaxon database (16S rRNA gene; Kim et al., 2012) and NCBI database (*gyrB* and *rpoD*; NCBI Resource Coordinators, 2016). The closest relatives based on 16S rRNA gene sequence similarity were identified against the EzTaxon database (Kim et al., 2012). The average nucleotide identity (ANI) and tetra-nucleotide signature correlation index (TETRA) were obtained using the online tool JSpeciesWS (Richter et al., 2016) and genome-to-genome distance calculator (GGDC) was obtained through the web service described in Meier-Kolthoff et al. (2013).

The 16S rRNA, *gyrB* and *rpoD* gene sequences were aligned using the online tool Clustal Omega (McWilliam et al., 2013) and the alignment was edited with BioEdit version 7.2.5 (Hall, 1999). Phylogenetic analyses were performed on the three genes separately and concatenated (Villesen, 2007) using MEGA version 6.0 (Tamura et al., 2013). Clustering using neighbour-joining (NJ, Saitou & Nei, 1987) and maximum-likelihood (ML, Felsenstein, 1981) methods was performed using the Kimura two-parameter model (Kimura, 1980). Clustering was additionally performed using the maximum parsimony method (MP, Fitch, 1971), and bootstrap values based on 1000 replications were calculated in all the phylogenetic analyses.

### **Morphological and phenotypic assays**

Conditions for optimal growth were assessed using the following medium (adapted from Ventosa et al., 1982): 5 g L<sup>-1</sup> proteose peptone no. 3 (Merck, USA), 10 g L<sup>-1</sup> yeast extract (Alfa Aesar, USA), 1 g L<sup>-1</sup> glucose (Merck, USA), and pH adjusted to 7.2 with KOH (2 M; JMGS, Portugal). NaCl tolerance was tested using concentrations of 0, 0.5, 1, 2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20, 25 and 30 % (w/v). At the optimum NaCl concentration for growth, the temperature range and optimum was tested by incubating at 4, 15, 20, 25, 30, 37, 40 and 45 °C. At the optimal growth temperature and NaCl

concentration for growth, the pH range and optimum was assessed for pH levels 4, 5, 6, 7, 8, 8.5, 9 and 10. For the pH assays, the medium was buffered adequately, using 4-Morpholineethanesulfonic acid (MES buffer from Sigma, USA; pH 4 to 5), piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES buffer from Panreac, Spain; pH 6 to 7), Tricine buffer (from Sigma, USA; pH 8 to 8.5) and 3-(Cyclohexylamino)-2-hydroxy-1-propanesulfonic acid (CAPSO buffer, from Sigma, USA; pH 9 to 10), all prepared at a concentration of 40 mM. The assays were conducted in 96-well plates, with 200  $\mu$ L of adequate medium and an inoculum of 5  $\mu$ L from a cell suspension with optical density at 590 nm ( $OD_{590}$ ) of 0.5, prepared in NaCl 0.9 % (w/v). The plates were incubated up to 7 days, and  $OD_{590}$  readings were registered at 0, 2, 4, 6, 8, 12, 24, 36, 48, 72, 120, 144 and 168 h. Growth curves were drafted and allowed determination of optimal growth conditions. A difference of 0.1 in  $OD_{590}$  from time 0 h was used as the cutoff for growth in these assays.

For the remaining tests, unless otherwise stated, strain growth was obtained on MA by incubating at 30 °C for 48 h. A Gram staining was performed using Merck's (Germany) Gram staining set, according to the manufacturer's instructions. Light microscopy allowed assessment of cell morphology, size and motility. Catalase activity was determined using the  $H_2O_2$ -reagent, and oxidase activity was evaluated using oxidase strips (both from Liofilchem, Italy). Oxygen metabolism was assessed using thioglycollate medium (Merck, Germany) supplemented with 7.5 % NaCl, and growth was monitored for 7 days. DNA hydrolysis was determined on DNase agar (HiMedia, India) supplemented with 0.1 % toluidine blue and with 7.5 % NaCl. Simmon's Citrate Agar (Oxoid, USA) supplemented with 7.5 % NaCl was used to determine citrate utilization. Kligler's Iron Agar (Merck, Germany) supplemented with 7.5 % NaCl was used to assess production of  $H_2S$ . Casein, starch, tween 20, cellulose, pectin (at pH 5 and 7) and xylan hydrolysis, phosphate solubilization, siderophore production, ACC deaminase activity and IAA production were assessed as described in Fidalgo et al. (2016). Miniaturized assays API 20NE, API ZYM and API 50CH (bioMérieux, France) were used according to the manufacturer's instructions, except for using a 4 % sea salts (Sigma, USA) solution to prepare cell suspensions. Closely related type strains of *Salinicola* species were assayed in the same conditions for comparative purposes.

The fatty acid profile and isoprenoid quinone determination was conducted as described in Proença et al. (2014) and performed in parallel with related *Salinicola* type strains. Growth for determination of isoprenoid quinones was obtained using Marine Broth (MB) and incubating at 30 °C for 72 h. The G+C content (mol%) was additionally determined by HPLC (Mesbah et al., 1989).

## Results and discussion

### Genomic analysis

The genome annotation performed with the RAST online tool (Aziz et al., 2008; Overbeek et al., 2014) allowed for the determination of several important metrics regarding the sequenced genomes, such as sequence size, information about contig number and length, genome sequencing coverage, and G+C content (Table 4.11). Genome size varied between 3.62 and 4.63 Mbp, and the number of contigs varied from 13 to 120. The smallest number of contigs that contains half the sequenced genome (L50) was relatively low overall (2 to 8, and 16 for *Salinicola peritrichatus* JCM 18795<sup>T</sup>), indicating that the Illumina Hiseq platform allows for very large read lengths, which is desirable for this type of work.

The G+C content obtained from genome sequence data differed from *wet lab* experiments in almost all instances (Table 4.11). This difference comes from the fact that, despite the importance of this value in new taxa description, experimentally determined G+C contents are prone to inaccuracies due to measurement error and heterogeneity in methodologies (Kim et al., 2015). These authors indicate that the difference between values obtained in *wet lab* experiments and genome sequence data may differ in more than 1 %, and this is also observed in our work. The higher accuracy that is inherent to using G+C contents obtained from genome data allows the use of the cutoff of 1 % in G+C content variation to validate the existence of a new species (Kim et al., 2015), which has already been established in online tools (Richter et al., 2016).

The G+C contents observed for the endophytic strains are well within the range determined for the *Halomonadaceae* family (52 to 68 mol%; Franzmann et al., 1988). Although there is no range established for the genus *Salinicola*, values reported for the species previously described are within a range of G+C content from 58.8 (*Salinicola salarius*; Kim et al., 2007) to 64.0 (*Salinicola rhizosphaerae*; Raju et al., 2016). Considering only the values obtained from genome sequence data – which is what will be done from this point on –, the values for G+C content of the endophytic strains are either within this range or slightly above (60.6 to 65.9, Table 4.11). Since there are only six members of the genus validly described thus far, and that the threshold of 10 % divergence is not crossed (Owen & Pitcher, 1985), the endophytic strains may be considered to belong to the genus *Salinicola*. The difference in G+C content that will allow validation of novel species is further explored in the pairwise comparison of the endophytic strains and type strains of *Salinicola*.



**Table 4.11** Information of sequenced genomes.

	CPA60 <sup>T</sup>	CPA62 <sup>T</sup>	CPA92 <sup>T</sup>	RZ23	CR45 <sup>T</sup>	CR50 <sup>T</sup>	CR57	<i>S. acroporae</i> LMG 28587 <sup>T</sup>	<i>S. halophilus</i> CECT 5903 <sup>T</sup>	<i>S. peritrichatus</i> JCM 18795 <sup>T</sup>	<i>S. salarius</i> DSM 18044 <sup>T</sup>	<i>S. socius</i> DSM 19940 <sup>T</sup>
Sequence size (bp) <sup>Δ</sup>	3996664	3875167	4360898	4375234	3904690	4420691	4302657	4377847	3620402	4628485	3888731	4095312
Number of contigs <sup>Δ</sup>	13	36	32	49	30	30	35	36	17	120	50	45
Shortest contig size <sup>Δ</sup>	171	149	145	125	159	342	133	137	239	174	571	134
Median sequence size <sup>Δ</sup>	31101	3158	32984	4550	3286	50400	66535	37715	36215	20551	45749	2324
Longest contig size <sup>Δ</sup>	1575882	958676	902169	651083	786467	811871	593809	734116	1354622	337132	477531	1220288
N50 value <sup>Δ</sup>	938409	480424	579807	280645	467686	393965	276347	303762	659543	74359	155119	482332
L50 value <sup>Δ</sup>	2	3	4	6	3	4	5	4	2	16	8	3
Coverage <sup>**</sup>	613X	536X	351X	377X	374X	347X	216X	389X	376X	216X	305X	393X
G+C content (mol%) <sup>Δ</sup>	60.6	63.5	65.8	65.9	62.2	64.3	63.9	63.6	64.8	62.4	62.5	62.2
G+C content (mol%) from other sources	66.2*	66.4*	67.7*	nd	66.3*	66.9*	66.6*	63.6†	63.6‡	59.6§	58.8	63.0#

Values given correspond to those obtained after splitting into scaffolds. <sup>Δ</sup>, data obtained from the RAST online tool; <sup>\*\*</sup> Coverage was calculated as  $L \cdot N / G$ , where L is the length of the reads, N the read number, and G the estimated genome length given by RAST; nd, not determined; \* data from HPLC G+C content (mol%) determination in the present work; †, data from Lepcha et al. (2015); ‡, data from Aguilera et al. (2007); §, data from Huo et al. (2013); ||, data from Kim et al. (2007); #, data from Anan'ina et al. (2007).

The full-length 16S rRNA gene sequences (1528 bp) of the endophytic and related *Salinicola* strains were obtained from RAST annotation. The sequences from the endophytic strains were then compared to the EzTaxon database (Kim et al., 2012) to obtain a 16S rRNA gene-based identification. The results are listed in Table 4.12 and clearly place all endophytic strains in the genus *Salinicola*. The lowest similarity percentage observed is higher than the cutoff for delimitation of a novel species based on 16S rRNA gene sequence (97 %, Stackebrandt & Goebel, 1994; Wayne et al., 1987), however, this gene sequence is not the most adequate to resolve closely related species in the family *Halomonadaceae* (de la Haba et al., 2010a).

To overcome this shortcoming of the 16S rRNA gene sequence-based approach, a genome sequence-based pairwise comparison was performed, in which the value for average nucleotide identity (ANI) of pairs of strains is obtained. The ANI value was calculated using the parameters for ANIb (BLAST-based ANI), ANIm (MUMmer-based ANI), in addition to the TETRA parameter (Richter & Rosselló-Móra, 2009) with the online tool JSpeciesSW (Richter et al., 2016). According to the considerations of the authors of the parameters and recent evaluations (Figueras et al., 2014), ANI > 95-96 % indicates that strains belong to the same species.

A traditional methodology to assess if a pair of strains belong to the same species is to perform DNA-DNA hybridization (DDH), however this approach is prone to experimental error and has been replaced by genome sequence-based approaches (Meier-Kolthoff et al., 2013). These authors proposed the calculation of *in silico* DDH (*isDDH*) by genome blast distance phylogeny (GBDP) as an alternative to the traditional DDH method. This *isDDH* value was then calculated for the endophytic strains against the closely related type *Salinicola* strains.

The G+C content difference was also calculated using the tool proposed by Meier-Kolthoff et al. (2013) and the pairwise difference of G+C content is also used as a parameter to validate the presence of a new species. A difference in G+C content over 1 % rejects the hypothesis that the pair of strains belongs to the same species (Meier-Kolthoff et al., 2014).

**Table 4.12** Length of 16S rRNA gene sequences obtained from RAST annotation and best matches observed by comparison against the EzTaxon database (Kim et al., 2012).

	CPA60 <sup>T</sup>	CPA62 <sup>T</sup>	CPA92 <sup>T</sup>	RZ23	CR45 <sup>T</sup>	CR50 <sup>T</sup>	CR57
Length (bp)	1529	1529	1528	1528	1529	1528	1528
Best matches (% similarity)	<i>S. socius</i> SMB35 <sup>T</sup> (99.09 %)	<i>S. halophilus</i> CG 4.1 <sup>T</sup> (98.97 %)	<i>S. peritrichatus</i> DY22 <sup>T</sup> (98.29 %)	<i>S. peritrichatus</i> DY22 <sup>T</sup> (98.08 %)	<i>S. salarius</i> KCTC 12664 <sup>T</sup> (99.38 %)	<i>S. acroporae</i> S4-41 <sup>T</sup> (99.50 %)	<i>S. acroporae</i> S4-41 <sup>T</sup> (99.50 %)
	<i>S. salarius</i> KCTC 12664 <sup>T</sup> (99.04 %)	<i>S. acroporae</i> S4-41 <sup>T</sup> (98.50 %)	<i>S. acroporae</i> S4-41 <sup>T</sup> (97.79 %)	<i>S. acroporae</i> S4-41 <sup>T</sup> (97.57 %)	<i>S. socius</i> SMB35 <sup>T</sup> (98.74 %)	<i>S. salarius</i> KCTC 12664 <sup>T</sup> (99.25 %)	<i>S. salarius</i> KCTC 12664 <sup>T</sup> (99.25 %)
	<i>S. zeshunii</i> N4 <sup>T</sup> (98.43 %)*	<i>S. zeshunii</i> N4 <sup>T</sup> (98.50 %)*	<i>S. salarius</i> KCTC 12664 <sup>T</sup> (97.67 %)	<i>S. salarius</i> KCTC 12664 <sup>T</sup> (97.47 %)	<i>S. acroporae</i> S4-41 <sup>T</sup> (98.50 %)	<i>S. socius</i> SMB35 <sup>T</sup> (98.95 %)	<i>S. socius</i> SMB35 <sup>T</sup> (98.95 %)
	<i>S. acroporae</i> S4-41 <sup>T</sup> (98.22 %)	<i>S. salarius</i> KCTC 12664 <sup>T</sup> (98.36 %)	<i>S. rhizosphaerae</i> MSSRFH1 <sup>T</sup> (97.50 %)	<i>S. rhizosphaerae</i> MSSRFH1 <sup>T</sup> (97.36 %)	<i>S. zeshunii</i> N4 <sup>T</sup> (98.29 %)*	<i>S. halophilus</i> CG 4.1 <sup>T</sup> (98.08 %)	<i>S. halophilus</i> CG 4.1 <sup>T</sup> (98.08 %)

\* *Salinicola zeshunii* is an effectively published species (Cao et al., 2013) but not validly published thus far.

The results of the genome sequence-based pairwise comparisons are given in Table 4.13 and show that the threshold of ANI value (> 95-96 %) for species delineation is not surpassed for the endophytic strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup> and CR50<sup>T</sup>. For these strains, some TETRA values are over the threshold for species boundary (> 0.99, Richter & Rosselló-Móra, 2009), however, this parameter should only be considered when it is in agreement with ANI values greater than 95-96 % (Richter & Rosselló-Móra, 2009).

The *isDDH* parameter confirms these five endophytic strains represent different species from the currently described in the genus *Salinicola*. The difference in values of G+C content between strains also supports our claim that the endophytic strains represent novel genomic species of *Salinicola*, as the value of 1 % in difference is surpassed in most comparisons. According to these results, the endophytic strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup> and CR50<sup>T</sup> represent novel *Salinicola* species. The endophytic strains RZ23 and CR57, however, present parameters for ANIb, ANim, TETRA, *isDDH*, and differences in G+C content that support that they are members of the same species as CPA92<sup>T</sup> and CR50<sup>T</sup>, respectively.

### Phylogenetic analysis

In order to perform the MLSA approach, required for the description of new taxa in the *Halomonadaceae* family (Arahal et al., 2007; de la Haba et al., 2012), the full sequences of the 16S rRNA gene, *gyrB* and *rpoD* genes were retrieved from the annotated genomes of the endophytic strains and the related *Salinicola* type species. Additional sequences were used, namely from the recently published *Salinicola rhizosphaerae* MSSRFH1<sup>T</sup> and *Halomonas elongata* DSM 2581<sup>T</sup>, and the accession numbers used are listed in Table 4.14.

Phylogenetic trees were built with NJ, ML and MP clustering methods, for each gene separately (Figure 4.6), and for the concatenated sequences of the three genes (Figure 4.7).

**Table 4.13** Pairwise comparison of genome sequence-based information of the endophytic strains and related type strains of the genus *Salinicola*.

		CPA60 <sup>T</sup>	CPA62 <sup>T</sup>	CPA92 <sup>T</sup>	RZ23	CR45 <sup>T</sup>	CR50 <sup>T</sup>	CR57	<i>S. acroporae</i> LMG 28587 <sup>T</sup>	<i>S. halophilus</i> CECT 5903 <sup>T</sup>	<i>S. peritrichatus</i> JCM 18795 <sup>T</sup>	<i>S. salarius</i> DSM 18044 <sup>T</sup>	<i>S. socius</i> DSM 19940 <sup>T</sup>
CPA60 <sup>T</sup>	ANib	–	76.53	77.87	77.87	86.94	81.56	81.60	81.44	76.82	79.47	82.82	83.19
	ANIm	–	84.54	84.84	84.83	88.52	85.68	85.67	85.64	84.48	85.11	86.1	86.28
	TETRA	–	0.9576	0.906	0.906	0.987	0.948	0.954	0.962	0.929	0.971	0.972	<u>0.991</u>
	<i>isDDH</i>	–	21.8	23.1	23.0	32.5	24.7	24.5	24.6	22.1	23.3	26.2	26.4
	Prob ≥ 70 %	–	0	0	0	0.28	0.01	0.01	0.01	0	0	0.02	0.02
	G+C dif. (%)	–	<b>2.95</b>	<b>5.22</b>	<b>5.34</b>	<b>1.57</b>	<b>3.7</b>	<b>3.33</b>	<b>3.05</b>	<b>4.24</b>	<b>1.84</b>	<b>1.91</b>	<b>1.59</b>
CPA62 <sup>T</sup>	ANib	76.64	–	77.94	77.97	77.35	77.31	77.20	77.11	84.64	77.7	77.41	77.15
	ANIm	84.54	–	84.91	84.93	84.67	84.72	84.74	84.64	87.67	84.73	84.76	84.57
	TETRA	0.957	–	0.948	0.949	0.971	0.970	0.971	0.971	0.986	0.960	0.960	0.972
	<i>isDDH</i>	21.8	–	22.7	22.7	22.1	22.4	22.2	22.1	28.6	22.4	22.4	22
	Prob ≥ 70 %	0	–	0	0	0	0	0	0	0.06	0	0	0
	G+C dif. (%)	<b>2.95</b>	–	<b>2.27</b>	<b>2.39</b>	<b>1.38</b>	0.75	0.38	0.1	<b>1.29</b>	<b>1.11</b>	<b>1.04</b>	<b>1.36</b>
CPA92 <sup>T</sup>	ANib	77.73	77.65	–	<u>98.99</u>	78.88	80.74	80.53	79.87	78.35	79.6	78.85	78.68
	ANIm	84.83	84.91	–	<u>99.10</u>	85.31	87.02	86.86	86.46	85.15	85.55	85.36	85.14
	TETRA	0.906	0.948	–	<u>0.999</u>	0.950	0.970	0.967	0.959	0.952	0.952	0.948	0.939
	<i>isDDH</i>	23.1	22.7	–	<u>92.2</u>	23.7	25.7	25.3	24.7	23.5	24.3	23.4	23.4
	Prob ≥ 70 %	0	0	–	96.47	0	0.01	0.01	0.01	0	0.01	0	0
	G+C dif. (%)	<b>5.22</b>	<b>2.27</b>	–	0.12	<b>3.65</b>	<b>1.52</b>	<b>1.9</b>	<b>2.17</b>	0.98	<b>3.38</b>	<b>3.31</b>	<b>3.63</b>
RZ23	ANib	77.87	77.97	<u>98.99</u>	–	78.99	80.63	80.46	79.78	78.87	79.60	79.09	78.82
	ANIm	84.83	84.93	<u>99.10</u>	–	85.32	87.00	86.81	86.43	85.25	85.52	85.37	85.13
	TETRA	0.906	0.949	<u>0.999</u>	–	0.950	0.970	0.967	0.959	0.953	0.951	0.947	0.940

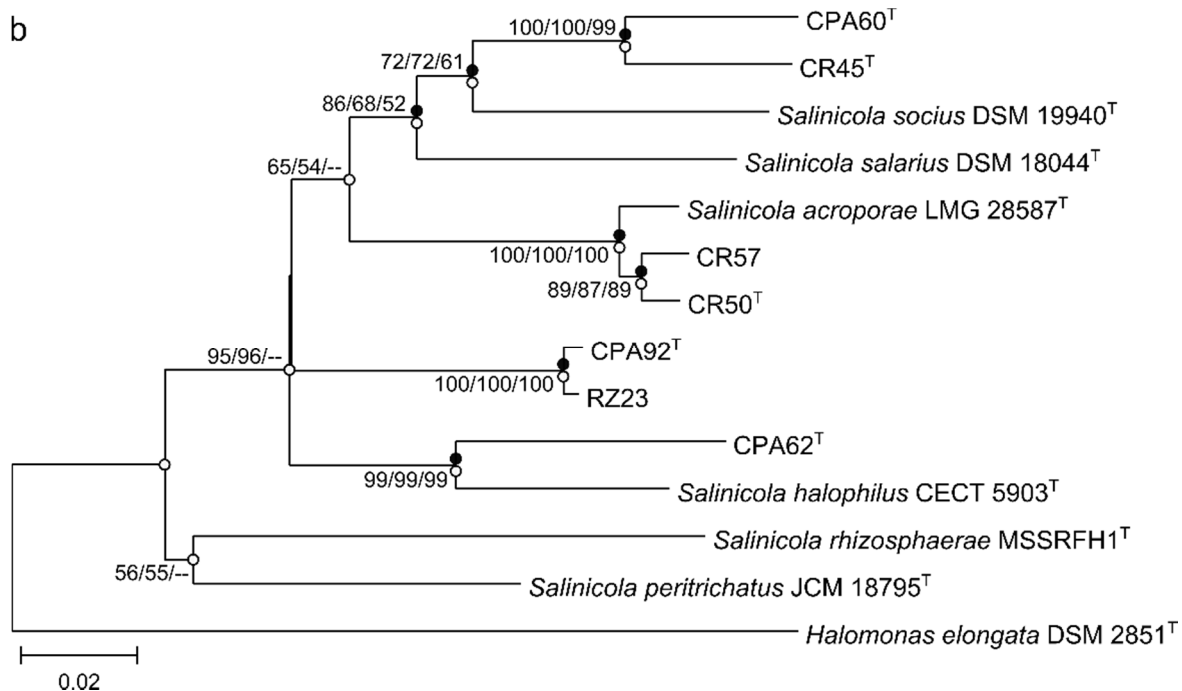
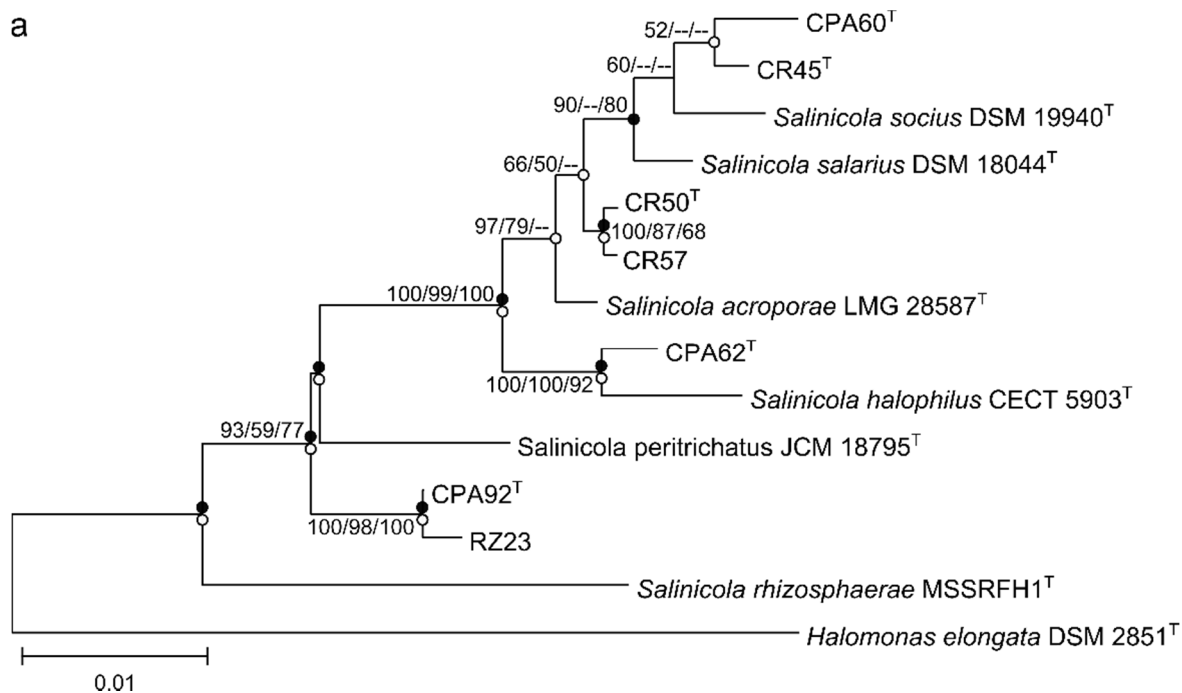
	<i>isDDH</i>	23.0	22.7	<u>92.2</u>	–	23.6	25.8	25.4	24.8	23.6	24.2	23.5	23.3
	Prob ≥ 70 %	0	0	96.47	–	0	0.01	0.01	0.01	0	0.01	0	0
	G+C dif. (%)	<b>5.34</b>	<b>2.39</b>	0.12	–	<b>3.77</b>	<b>1.64</b>	<b>2.01</b>	<b>2.29</b>	<b>1.1</b>	<b>3.5</b>	<b>3.43</b>	<b>3.75</b>
CR45 <sup>T</sup>	ANIb	86.96	77.11	79.02	78.99	–	82.64	82.57	82.29	77.68	80.41	83.62	84.98
	ANIm	88.52	84.67	85.3	85.32	–	86.23	86.22	86.07	84.73	85.46	86.59	87.76
	TETRA	0.987	0.971	0.950	0.950	–	0.977	0.980	0.984	0.958	0.988	<u>0.991</u>	<u>0.997</u>
	<i>isDDH</i>	32.5	22.1	23.7	23.6	–	26	25.9	25.5	22.7	24.4	27.2	29.4
	Prob ≥ 70 %	0.28	0	0	0	–	0.02	0.01	0.01	0	0.01	0.03	0.08
	G+C dif. (%)	<b>1.57</b>	<b>1.38</b>	<b>3.65</b>	<b>3.77</b>	–	<b>2.13</b>	<b>1.75</b>	<b>1.48</b>	<b>2.67</b>	0.26	0.34	0.02
CR50 <sup>T</sup>	ANIb	81.39	76.95	80.67	80.63	82.26	–	<u>96.40</u>	93.77	77.64	80.08	82.97	82.94
	ANIm	85.68	84.73	87.03	87.00	86.23	–	<u>96.96</u>	94.43	84.94	85.53	86.58	86.48
	TETRA	0.948	0.970	0.970	0.970	0.978	–	<u>0.999</u>	<u>0.997</u>	0.963	0.980	0.981	0.977
	<i>isDDH</i>	24.7	22.4	25.7	25.8	26	–	<u>72.9</u>	56.2	22.6	23.9	26.8	26.6
	Prob ≥ 70 %	0.01	0	0.01	0.01	0.02	–	83.02	39.09	0	0	0.02	0.02
	G+C dif. (%)	<b>3.7</b>	0.75	<b>1.52</b>	<b>1.64</b>	<b>2.13</b>	–	0.37	0.65	0.55	<b>1.86</b>	<b>1.79</b>	<b>2.11</b>
CR57	ANIb	81.60	77.20	80.53	80.58	82.57	<u>96.40</u>	–	93.56	77.92	79.89	83.13	83.11
	ANIm	85.67	84.74	86.86	86.81	86.22	<u>96.96</u>	–	94.37	84.90	85.49	86.46	86.52
	TETRA	0.954	0.971	0.967	0.967	0.980	<u>0.999</u>	–	<u>0.999</u>	0.961	0.982	0.984	0.980
	<i>isDDH</i>	24.5	22.2	25.3	25.4	25.9	<u>72.9</u>	–	55.6	22.4	23.9	26.6	26.4
	Prob ≥ 70 %	0.01	0	0.01	0.01	0.01	83.02	–	37.06	0	0	0.02	0.02
	G+C dif. (%)	<b>3.33</b>	0.38	<b>1.9</b>	<b>2.01</b>	<b>1.75</b>	0.37	–	0.28	0.92	<b>1.49</b>	<b>1.41</b>	<b>1.74</b>

ANIb, BLAST-based ANI; ANIm, MUMmer-based ANI; TETRA, tetra-nucleotide signature correlation index; *isDDH*, *in silico* DNA-DNA hybridization; Prob ≥ 70 %, probability that the *isDDH* calculation is 70 % or higher; G+C dif., difference in G+C content (mol%). Underlined values for ANIb, ANIm, TETRA and *isDDH* parameters support that the strains in the pairwise comparison belong to the same species; bold values in G+C dif. support that the strains in the pairwise comparison belong to different species.

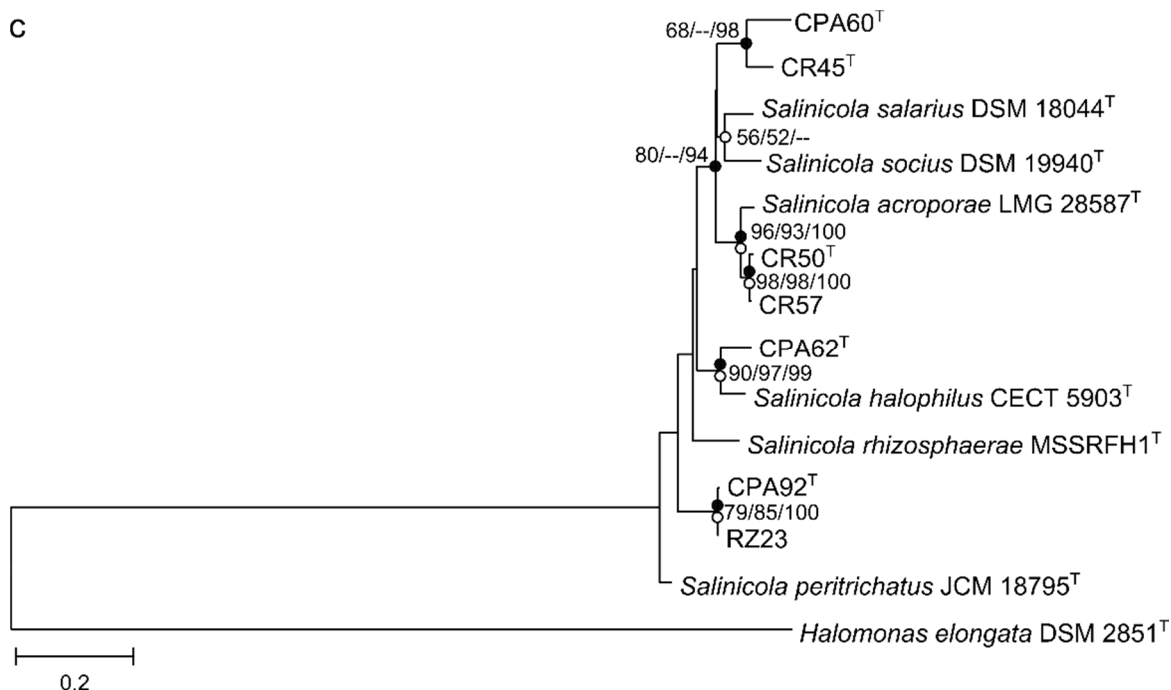
**Table 4.14** Length and origin of sequences used in the MLSA approach.

	CPA60 <sup>†</sup>	CPA62 <sup>†</sup>	CPA92 <sup>†</sup>	RZ23	CR45 <sup>†</sup>	CR50 <sup>†</sup>	CR57	<i>S. acroporae</i> LMG 28587 <sup>†</sup>	<i>S. halophilus</i> CECT 5903 <sup>†</sup>	<i>S. peritrichatus</i> JCM 18795 <sup>†</sup>	<i>S. salarius</i> DSM 18044 <sup>†</sup>	<i>S. socius</i> DSM 19940 <sup>†</sup>	<i>S. rhizosphaerae</i> MSSRFH1 <sup>†</sup>	<i>Halomonas</i> <i>elongata</i> DSM 2581 <sup>†</sup>
16S rRNA gene*	1529	1528	1528	1528	1529	1528	1528	1528	1528	1528	1529	1529	1442‡ (KF898345)	1460‡ (FN869568)
<i>gyrB</i> <sup>†</sup>	2415	2415	2415	2415	2415	2415	2415	2415	2415	2418	2415	2415	1050§ (FN869568)	2421§ (KR996750)
<i>rpoD</i> <sup>†</sup>	1848	1848	1863	1863	1848	1848	1848	1848	1848	1854	1848	1845	1121§ (FN869568)	1848§ (KJ495727)

\*, sequences extracted with RNAmmer; †, sequences extracted from RAST annotation; ‡, sequences obtained from EzTaxon database; §, sequences obtained from NCBI database. Accession numbers are given for sequences from databases.

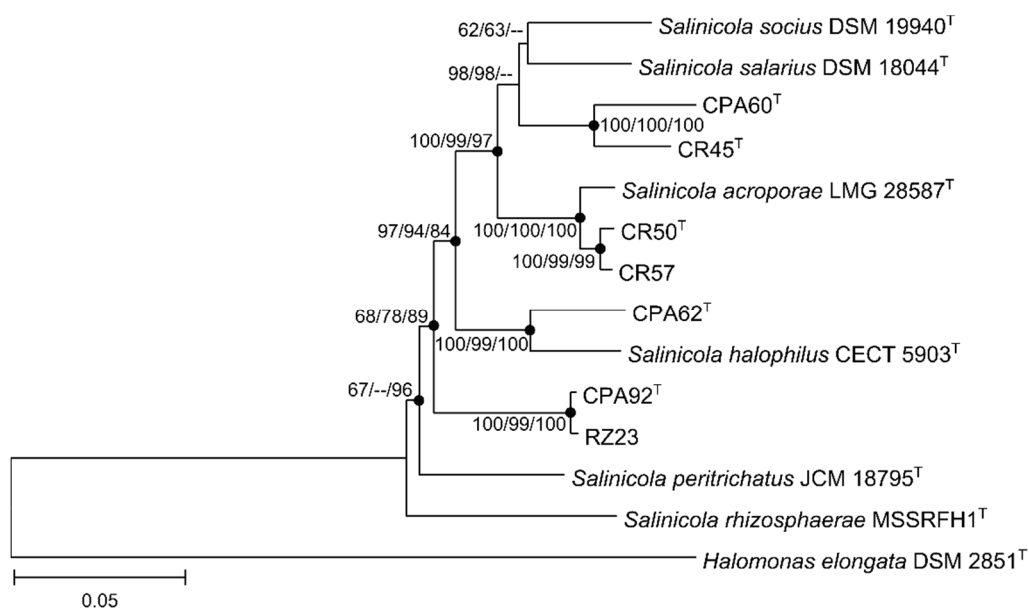






**Figure 4.6** Neighbour-joining (NJ) phylogenetic trees showing the phylogenetic positions of strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, RZ23, CR45<sup>T</sup>, CR50<sup>T</sup>, CR57 and representatives of other related taxa based on 16S rRNA gene (a), *gyrB* (b) and *rpoD* (c) sequences. Filled circles indicate nodes that were also recovered in maximum parsimony (MP) trees, and empty circles indicate nodes also recovered from maximum-likelihood (ML) trees, based on the same sequences. Bootstrap values (expressed as percentages of 1000 replications)  $\geq 50\%$  are shown at branch points for NJ/ML/MP trees. *Halomonas elongata* DSM 2581<sup>T</sup> was used as outgroup. Bar indicates nucleotide substitution rate ( $K_{nuc}$ ) units.

The phylogenetic trees obtained for each separate gene show that the congruence among clustering methods is not complete, since several of the branches observed in the NJ tree are not mirrored in the ML and MP trees. This was especially true for the *rpoD* gene-based phylogenetic tree (Figure 4.6c). This result highlights the importance of using a MLSA approach. The MLSA tree shows higher, albeit not complete, congruence among the three clustering methods (Figure 4.7). This indicates that the observed topology and phylogenetic positions of the endophytic strains are robust and stable, supporting our claim that the endophytic strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup> and CR50<sup>T</sup> are, in fact, new species of the *Salinicola* genus, and that strains RZ23 and CR57 belong to the same species as CPA92<sup>T</sup> and CR50<sup>T</sup>, respectively.



**Figure 4.7** Neighbor-joining (NJ) phylogenetic tree showing the phylogenetic positions of strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, RZ23, CR45<sup>T</sup>, CR50<sup>T</sup>, CR57 and representatives of other related taxa based on the concatenated sequences of the 16S rRNA, *gyrB* and *rpoD* genes. The topology of the maximum-likelihood (ML) tree was the same as for the NJ tree. Filled circles indicate nodes that were also recovered in maximum parsimony (MP) trees, based on the same sequences. Bootstrap values (expressed as percentages of 1000 replications)  $\geq 50\%$  are shown at branch points for NJ/ML/MP trees. *Halomonas elongata* DSM 2581<sup>T</sup> was used as outgroup. Bar indicates nucleotide substitution rate ( $K_{nuc}$ ) units.

### Morphological and phenotypic analysis

The optimal conditions for growth were determined and are detailed in the species description for each type strain. All strains are moderately halophilic and mesophilic, which is in accordance with what has been observed for validly described *Salinicola* species. All strains are Gram-negative, catalase- and oxidase-positive motile rods, facultative anaerobes that do not produce H<sub>2</sub>S. Citrate degradation was observed for all strains (weak in the case of CPA62<sup>T</sup>), as was Tween 20 hydrolysis, phosphate solubilization and IAA production. Casein was not degraded by any of the endophytic strains.

In the API 20NE strip, all strains were negative for reduction of nitrates to nitrites, indole production, fermentation of D-glucose, arginine dihydrolase, hydrolysis of gelatin (protease), and assimilation of adipic acid and phenylacetic acid. In the API ZYM strip, all strains were positive for acid and alkaline phosphatase, leucine and valine arylamidase, and negative for lipase (C14), cystine

arylamidase, trypsin,  $\alpha$ -chymotrypsin,  $\alpha$ -galactosidase,  $\beta$ -galactosidase,  $\beta$ -glucuronidase, *N*-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -mannosidase and  $\alpha$ -fucosidase. In the API 50CH strip, all strains were positive for D-galactose, D-glucose, D-ribose, D-xylose and L-arabinose assimilation, and negative for amidon (starch), amygdalin, arbutin, D-adonitol, D-arabitol, D-melezitose, D-raffinose, gentiobiose, glycogen, inositol, inulin, L-arabitol, L-fucose, L-xylose, methyl- $\alpha$ -D-glucopyranoside, methyl- $\alpha$ -D-mannopyranoside, methyl- $\beta$ -D-xylopyranoside, *N*-acetylglucosamine, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, salicin and xylitol. The variable results for the remaining assays are detailed in the species description for each novel *Salinicola* species.

The Q-9 was the major respiratory quinone for all strains, which is in agreement with what has been described for the genus. The major fatty acids overall were C<sub>16:0</sub> and the summed feature containing C<sub>18:1</sub>  $\omega$ 7c and/or C<sub>18:1</sub>  $\omega$ 6c. The cyclo-C<sub>19:0</sub>  $\omega$ 8c is also a major fatty acid but only for strains CR45<sup>T</sup> and CR50<sup>T</sup>. The complete fatty acid profiles of the endophytic strains and reference strains are given in Table 4.15. According to the description of the genus *Salinicola*, C<sub>16:1</sub>  $\omega$ 7c, C<sub>16:0</sub>, C<sub>18:1</sub>  $\omega$ 7c and cyclo-C<sub>19:0</sub> are the major fatty acids (Anan'ina et al., 2007). The unsaturated fatty acid C<sub>16:1</sub>  $\omega$ 7c was listed as a major fatty acid by Anan'ina et al. (2007) when the genus was composed by a single species, despite the fact that this fatty acid comprised only 8.6 % of the total fatty acids, and has not been considered a major fatty acid since that publication. The cyclo-C<sub>19:0</sub> described in Anan'ina et al. (2007) as a major fatty acid has not had the same designation in other papers, where it has been referenced as cyclo-C<sub>19:0</sub>  $\omega$ 8c, and has been, in fact, observed in high percentages. The C<sub>16:0</sub> and C<sub>18:1</sub>  $\omega$ 7c fatty acids have been observed as major fatty acids in every *Salinicola* species validly published so far, and this is also observed with the endophytic strains, where the latter is listed under the summed feature containing C<sub>18:1</sub>  $\omega$ 7c and/or C<sub>18:1</sub>  $\omega$ 6c. A more appropriate characteristic list of main fatty acids is given in the emended description of the genus, based not only on our results, but also on those published since the last amendment.

Considering the genomic, phylogenetic and biochemical and physiological characteristics, it is clear the CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup>, CR50<sup>T</sup> strains are affiliated with the genus *Salinicola*. These strains are, nevertheless, distinguishable from other validly described *Salinicola* species, and differentiating traits are stated in Table 4.16. Accordingly, CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup>, CR50<sup>T</sup> represent novel species of the genus *Salinicola*, and the names *Salinicola halimionae* sp. nov., *Salinicola aestuarina* sp. nov., *Salinicola endophytica* sp. nov., *Salinicola halophytica* sp. nov. and *Salinicola lusitana* sp. nov., respectively, are proposed.

**Table 4.15** Complete fatty acid composition of the endophytic strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup> and CR50<sup>T</sup>, and related type strains.

	CPA60 <sup>T</sup>	CPA62 <sup>T</sup>	CPA92 <sup>T</sup>	CR45 <sup>T</sup>	CR50 <sup>T</sup>	<i>S. acroporae</i> S4-41 <sup>T</sup> †	<i>S. halophilus</i> CECT 5903 <sup>T</sup>	<i>S. halophilus</i> CG4.1 <sup>T</sup> ‡	<i>S. peritrichatus</i> JCM 18795 <sup>T</sup>	<i>S. peritrichatus</i> DY22 <sup>T</sup> §	<i>S. salarius</i> DSM 18044 <sup>T</sup>	<i>S. salarius</i> M27 <sup>T</sup>	<i>S. socius</i> DSM 19940 <sup>T</sup>	<i>S. socius</i> SMB35 <sup>T</sup> #	<i>S. rhizosphaerae</i> MSSRFH1 <sup>T</sup> ¶
<b>Saturated</b>															
C <sub>10:0</sub>	tr	1.2	1.1	1.3	tr	–	1.1	–	1.2	1.8	1.4	2.2	tr	–	2.4
C <sub>12:0</sub>	2.9	3.7	1.5	3.6	2.0	4.4	3.0	tr	1.9	2.4	3.1	3.0	2.6	–	4.4
C <sub>12:0</sub> aldehyde	–	–	–	–	–	7.0	–	–	–	–	–	–	–	–	–
C <sub>14:0</sub>	tr	–	tr	tr	tr	8.6	tr	tr	tr	tr	–	–	–	–	1.0
C <sub>16:0</sub>	22.2	22.5	28.1	21.9	27.6	35.4	25.9	32.6	29.9	33.7	26.8	20.4	21.5	19.3	22.7
C <sub>17:0</sub>	–	tr	1.2	1.4	tr	1.1	tr	–	tr	tr	tr	–	tr	–	–
C <sub>18:0</sub>	9.9	8.1	8.4	6.7	3.6	tr	2.6	tr	2.1	2.6	2.4	1.3	3.6	–	1.5
<b>Unsaturated</b>															
C <sub>16:1</sub> ω7c	–	–	–	–	–	–	–	–	–	–	–	–	–	8.6	–
C <sub>18:1</sub> ω7c	–	–	–	–	–	1.7	–	43.0	–	7.0	–	33.6	–	46.7	33.7
C <sub>18:1</sub> ω7c 11-methyl	–	tr	–	tr	1.0	tr	1.2	tr	1.6	1.0	1.2	–	1.2	–	–
C <sub>20:2</sub> ω6,9c	–	tr	–	tr	tr	tr	tr	–	1.2	tr	tr	–	tr	–	–
<b>Branched-chain</b>															
anteiso-C <sub>15:0</sub>	–	–	–	2.2	1.9	tr	–	–	–	–	–	–	–	–	–
anteiso-C <sub>17:0</sub>	–	–	–	1.3	tr	tr	–	–	–	–	–	–	–	–	–
iso-C <sub>16:0</sub>	–	tr	–	tr	tr	–	tr	–	–	tr	tr	–	1.0	–	–
iso-C <sub>18:0</sub> 10(11)-methyl	–	–	–	–	1.1	–	–	–	–	–	–	–	–	–	–

Cyclic															
cyclo-C <sub>17:0</sub>	–	1.3	–	3.3	2.3	24.4	3.5	1.7	5.2	4.0	4.0	2.9	1.7	–	1.9
cyclo-C <sub>19:0</sub> ω8c	1.3	12.4	8.1	25.5	30.6	6.1	23.3	11.1	42.9	29.0	26.3	18.3	21.9	7.4*	14.0
Hydroxyl															
C <sub>12:0</sub> 2-OH	tr	1.0	tr	tr	1.9	–	1.3	–	1.7	3.5	1.4	–	1.9	–	tr
C <sub>12:0</sub> 3-OH	6.5	7.4	6.3	6.0	6.1	tr	8.1	tr	5.4	10.1	8.4	8.3	7.8	–	5.0
C <sub>14:0</sub> 2-OH	–	–	1.7	tr	–	–	–	–	–	tr	–	–	–	–	–
C <sub>14:0</sub> 3-OH/iso-C <sub>16:1</sub> I	–	–	–	–	–	7.0	–	–	–	–	–	–	–	–	–
Summed feature															
2	tr	2.3	2.1	2.8	2.3	–	tr	–	–	tr	tr	–	1.9	–	–
3	10.5	5.4	4.6	1.6	tr	–	5.0	10.1	1.0	1.7	3.7	10.0	3.8	–	–
8	44.1	33.2	34.6	18.0	12.2	–	20.9	–	4.4	–	18.5	–	28.7	–	–
Sa	–	–	–	–	–	2.5	–	–	–	–	–	–	–	–	–
Sr 1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	4.3
Sr 2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	5.3

Summed feature 2 contains C<sub>12:0</sub> aldehyde, C<sub>14:0</sub> 3-OH and/or iso-C<sub>16:1</sub> and/or Unknown Equivalent Chain Length (ECL) 10.927; summed feature 3 contains C<sub>16:1</sub> ω7c and/or iso-C<sub>15:0</sub> 2-OH; summed feature 8 contains C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c; summed feature Sa (*Salinicola acroporae*, from Lepcha et al., 2015) contains C<sub>16:1</sub> ω7c and/or C<sub>16:1</sub> ω6c; summed feature Sr 1 (*Salinicola rhizosphaerae*, from Raju et al., 2016) contains C<sub>14:0</sub> 3-OH/iso-C<sub>16:1</sub> I; summed feature Sr 2 (Raju et al., 2016) contains C<sub>16:1</sub> ω7c/C<sub>16:1</sub> ω6c. \* Referred as cyclo-C<sub>19:0</sub> ω8c in Anan'ina et al. (2007); –, not detected; tr, trace (< 1 %); †, data from Lepcha et al. (2015); ‡, data from Aguilera et al. (2007); §, data from Huo et al. (2013); ||, data from Kim et al. (2007); #, data from Anan'ina et al. (2007); ¥, data from Raju et al. (2016).

**Table 4.16** Differential and phenotypic characteristics of strains CPA60<sup>T</sup>, CPA62<sup>T</sup>, CPA92<sup>T</sup>, CR45<sup>T</sup> and CR50<sup>T</sup>, and related type strains.

Characteristic	CPA60 <sup>T</sup>	CPA62 <sup>T</sup>	CPA92 <sup>T</sup>	CR45 <sup>T</sup>	CR50 <sup>T</sup>	<i>S. acroporae</i> S4-41 <sup>T</sup> ‡	<i>S. halophilus</i> CECT 5903 <sup>T</sup>	<i>S. peritrichatus</i> JCM 18795 <sup>T</sup>	<i>S. salarius</i> DSM 18044 <sup>T</sup>	<i>S. socius</i> DSM 19940 <sup>T</sup>	<i>S. rhizosphaerae</i> MSSRFH1 <sup>T</sup> #
Relation to oxygen	facultative anaerobe	facultative anaerobe	facultative anaerobe	facultative anaerobe	facultative anaerobe	aerobic	aerobic*	aerobic†	aerobic‡	aerobic§	aerobic
Catalase	+	+	+	+	+	+	+	–†*	–‡*	–§*	–
Oxidase	+	+	+	+	+	+	–	+†*	+	+§*	–
Temperature range (°C)	15-40	15-40	4-45	4-40	15-45	15-45	15-45*	4-42†	10-45‡	15-40	20-37
pH range	4-10	4-10	4-10	5-10	4-10	4.5-9	5-9*	4.5-8.5†	5-10‡	5-9	4-9
NaCl (w/v) range	0.5-25 %	2.5-30 %	0-25 %	2.5-20 %	0-25 %	0-25 %	3-25 %*	0-20 %†	0-25 %‡	0.5-30 %§	0.5-30 %
ACC deaminase activity	+	–	+	+	+	nd	nd	nd	nd	nd	nd
Hydrolysis of:											
Casein	–	–	–	–	–	nd	–*	+†	+‡	nd	nd
Cellulose	+	–	+	–	+	nd	nd	nd	nd	nd	nd
Citrate	+	+	+	+	+	nd	+*	nd	nd	nd	–
DNA	+	–	w	w	+	+	–	–†	+‡	nd	nd
Pectin (pH 5.0)	+	nd	–	nd	–	nd	nd	nd	nd	nd	nd
Pectin (pH 7.0)	+	nd	–	+	+	nd	nd	nd	nd	nd	nd
Starch	+	–	+	–	+	–	+*	–†	–‡	+§	–
Tween 20	+	+	+	+	+	–	+*	+†	nd	nd	nd
Xylan	–	+	+	–	–	nd	nd	nd	nd	nd	nd
API 20NE results:											
Reduction of nitrates to nitrites	–	–	–	–	–	–	–**	–†*	+	–	nd
Fermentation of D-glucose	–	–	–	–	–	nd	–**	–†*	–	–	nd
Arginine dihydrolase	–	–	–	–	–	–	–	–†*	–	–	nd
Urease	–	–	+	–	–	nd	–	–	–	–	nd

Aesculin hydrolysis ( $\beta$ -glucosidase)	+	–	–	–	–	nd	–	–	–	–	nd
Para-nitrophenyl- $\beta$ -D- galactopyranose	+	–	–	–	–	–	–	–	–	–	nd
Assimilation of D-glucose, D-mannose	+	–	+	+	+	+	–	+	+	+	nd
Assimilation of L-arabinose	+	–	+	+	+	nd	–	+	+	–	nd
Assimilation of D-mannitol	+	–	+	+	+	+	w	+	w	+	nd
Assimilation of <i>N</i> -acetylglucosamine	–	–	+	–	+	nd	–	+	–	+	nd
Assimilation of D-maltose	–	–	+	–	+	nd	–	–†*	–	–	nd
Assimilation of potassium gluconate	+	–	+	+	+	+	–	–	+	–	nd
Assimilation of capric acid	–	–	+	–	+	nd	–	–	–	–	nd
Assimilation of adipic acid	–	–	–	–	–	–	–	–	+	–	nd
Assimilation of malic acid	+	–	+	–	+	+	–	w	w	–	nd
Assimilation of trisodium citrate	+	–	+	–	+	+	–	–†*	w	–	nd
Assimilation of phenylacetic acid	–	–	–	–	–	+	–	–	–	–	nd
API ZYM results:											
Esterase (C4)	+	+	+	–	+	+	w	–†*	–	+	nd
Esterase lipase (C8)	+	w	+	–	+	+	w	–†*	–	–	nd
Lipase (C14)	–	–	–	–	–	+	–	–	–	–	nd
Valine arylamidase	+	+	+	+	+	+	+	+	+‡*	+	nd
Naphthol-AS-BI- phosphohydrolase	+	w	+	w	+	+	w	+	w‡*	w	nd
$\alpha$ -glucosidase	+	+	+	–	+	–	w	+	–	–	nd
$\beta$ -glucosidase	+	–	–	–	–	–	–	–	–	–	nd
API 50CH results:											
D-arabinose	–	w	+	–	w	+	–	–†*	–	w	nd
D-arabitol	–	–	–	–	–	+	–	–	–‡*	–	nd

D-cellobiose	–	–	W	+	–	+	W**	–†*	–‡*	–	nd
D-fructose	+	W	+	+	W	+	W	–†*	–‡*	–§*	+
D-fucose	+	W	+	+	+	+	***	–†*	+	+	nd
D-lactose	–	–	W	W	–	+	W	–†*	–‡*	–	nd
D-lyxose	–	–	+	–	–	+	–	–†*	–‡*	–	nd
D-maltose	–	–	–	W	–	+	+	W†*	+	–§*	nd
D-mannitol	+	–	+	–	–	+	–	–	–‡*	–§*	+
D-mannose	+	+	+	+	W	+	–	–†*	–‡*	–	nd
D-melibiose	–	–	W	–	–	+	–	–†*	–‡*	–	nd
D-ribose, D-xylose	+	+	+	+	+	+	***	+	+	+	nd
D-saccharose	–	–	W	–	–	–	***	W†*	W‡*	–	nd
D-sorbitol	–	–	+	–	–	–	–	–	–	–§*	nd
D-tagatose, L-sorbose	–	–	W	–	–	–	–	–	–	–	nd
D-trehalose	+	–	–	–	–	+	***	W†*	+	–§*	–
D-turanose, dulcitol	–	–	+	–	–	–	–	–	–	–	nd
Erythritol	+	–	+	–	W	+	–	–†*	–‡*	W	nd
Aesculin ferric citrate	–	–	–	W	–	–	+	W†*	W‡*	–	nd
Glycerol	+	–	–	–	–	+	–	+	–‡*	–§*	nd
L-arabinose	+	+	+	+	+	+	***	+	+	W§*	nd
L-rhamnose	–	–	W	–	–	–	–	–†*	–‡*	–	nd
N-acetylglucosamine	–	–	–	–	–	–	+	W†*	W‡*	–	nd
Salicin	–	–	–	–	–	–	***	–	–	–	nd
DNA G+C content (mol%) =	60.6	63.5	65.8	62.2	64.3	63.6	64.8	62.4	62.5	62.2	64.0

¥, data from Lepcha et al. (2015); #, data from Raju et al. (2016); \*, data from Aguilera et al. (2007); \*\*, our result differed from those obtained by Aguilera et al. (2007); †, data from Huo et al. (2013); †\*, our result differed from that obtained in Huo et al. (2013); ‡, data from Kim et al. (2007); ‡\*, our result differed from that obtained by Kim et al. (2007); §, data from Anan'ina et al. (2007); §\*, our result differed from those obtained by Anan'ina et al. (2007); ||, data from de la Haba et al (2010b); =, data from genome data from this study except for *S. rhizosphaerae*; +, positive; –, negative; nd, not determined.



## **Emended description of the genus *Salinicola* Anan'ina et al. 2007, emend. de la Haba et al. 2010b**

The description of the genus *Salinicola* is as given by Anan'ina et al. (2007) and de la Haba et al. (2010b) with the following amendments: G+C content ranges from 58.8 to 66.9 mol%, oxidase, catalase and DNase activities are variable, temperature range and optimum is 4 to 45 and 25 to 37 °C, respectively. The major fatty acids observed are cyclo-C<sub>19:0</sub> ω8c, C<sub>16:0</sub> and C<sub>18:1</sub> ω7c. Hydrolysis of Tween 20, casein, DNA and aesculin is variable.

### **Description of *Salinicola halimionae* sp. nov. CPA60<sup>T</sup>**

*Salinicola halimionae* (N.L. n. *halimione* a botanical genus name; N.L. gen. n. *halimionae* of Halimione, isolated from *Halimione portulacoides*).

The CPA60<sup>T</sup> strain is a Gram-negative, positive for catalase and oxidase, facultative anaerobe, motile rod (1.81–2.96 μm L × 0.72–0.90 μm W). Colonies are whitish, and have round smooth edges, with 1–1.3 mm in diameter, after incubation in MA for 48 h at 30 °C. Salt is required for growth, which is observed at salinities of 0.5 to 25 % NaCl (w/v) with a salinity optimum observed at the 5 to 10 % NaCl (w/v) interval; at temperatures ranging from 15 to 40 °C, with optimum growth at 30 °C; and from pH 4.0 to 10.0, where the optimum is observed at pH 6.0. Production of H<sub>2</sub>S is not observed. Cellulose, DNA, pectin, starch, Tween 20 and citrate are hydrolyzed, however, casein and xylan are not. Phosphate solubilization, siderophore production and ACC deaminase activity are observed. IAA production is observed but low (4.7 μg mL<sup>-1</sup>). In the API 20NE strip, the strain is positive for β-glucosidase (aesculin), β-galactosidase (para-nitrophenyl-β-D-galactopyranose), assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, potassium gluconate, malic acid and trisodium citrate; and negative for reduction of nitrates to nitrites, indole production, fermentation of D-glucose, arginine dihydrolase, urease, protease (gelatin), assimilation of *N*-acetylglucosamine, D-maltose, capric acid, adipic acid and phenylacetic acid. In the API ZYM strip, the strain is positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-glucosidase and β-glucosidase; and negative for lipase (C14), cysteine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, *N*-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase. In API 50CH, the strain utilizes D-fructose, D-fucose, D-galactose, D-glucose, D-mannitol, D-mannose, D-ribose, D-trehalose, D-xylose, erythritol, glycerol and L-

arabinose. The strain does not utilize amidon, amygdalin, arbutin, D-adonitol, D-arabinose, D-arabitol, D-cellobiose, D-lactose, D-lyxose, D-maltose, D-melezitose, D-melibiose, D-raffinose, D-saccharose, D-sorbitol, D-tagatose, D-turanose, dulcitol, aesculin, gentiobiose, glycogen, inositol, inulin, L-arabitol, L-fucose, L-rhamnose, L-sorbose, L-xylose, methyl- $\alpha$ -D-glucopyranoside, methyl- $\alpha$ -D-mannopyranoside, methyl- $\beta$ -D-xylopyranoside, *N*-acetylglucosamine, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, salicin and xylitol. The main fatty acids are C<sub>16:0</sub> and summed feature 8, comprising C<sub>18:1</sub>  $\omega$ 7c and/or C<sub>18:1</sub>  $\omega$ 6c. The main respiratory quinone is Q-9.

The type strain, CPA60<sup>T</sup>, was isolated from the surface sterilized aboveground tissues of the halophyte *Halimione portulacoides*. The G+C content of the type strain, obtained from genomic data, is 60.6 mol%.

### **Description of *Salinicola aestuarina* sp. nov. CPA62<sup>T</sup>**

*Salinicola aestuarina* (L. n. *aestuarium*, a tidal marsh; L. suff. *-inus -a -um*, suffix used with the sense of belonging to; N.L. fem. adj. *aestuarina*, belonging to a marsh, referring to the location where the type strain was isolated).

The CPA62<sup>T</sup> strain is a Gram-negative, positive for catalase and oxidase, facultative anaerobe, motile rod (1.61–3.76  $\mu$ m L  $\times$  0.64–0.92  $\mu$ m W). Colonies are yellow, opaque but less so around the edges, with 1–1.5 mm in diameter, after incubation in MA for 48 h at 30 °C. Salt is required for growth, which is observed at salinities from 2.5 to 30 % NaCl (w/v) with a salinity optimum observed at the 10 to 12.5 % NaCl (w/v) interval; at temperatures ranging from 15 to 40 °C, with optimum growth at 30 °C; and from pH 4.0 to 10.0, where the optimum is observed at the interval of 6.0 to 7.0. Production of H<sub>2</sub>S is not observed. Tween 20, citrate and xylan are hydrolyzed, but, casein, cellulose, DNA and starch are not. No growth is observed in the medium used to detect pectin hydrolysis or siderophore production. Phosphate solubilization and IAA production (79.2  $\mu$ g mL<sup>-1</sup>) are observed, however, ACC deaminase activity is not. In the API 20NE strip, the strain is negative for all substrates. In the API ZYM strip, the strain is positive for alkaline phosphatase, esterase (C4), leucine arylamidase, valine arylamidase, acid phosphatase and  $\alpha$ -glucosidase. Weak reactions are observed for esterase lipase (C8) and naphthol-AS-BI-phosphohydrolase; and negative results are observed for lipase (C14), cysteine arylamidase, trypsin,  $\alpha$ -chymotrypsin,  $\alpha$ -galactosidase,  $\beta$ -galactosidase,  $\beta$ -glucuronidase,  $\beta$ -glucosidase, *N*-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -mannosidase and  $\alpha$ -

fucosidase. In API 50CH, the strain utilizes D-galactose, D-glucose, D-mannose, D-ribose, D-xylose and L-arabinose. Weak reactions are observed for D-arabinose, D-fructose and D-fucose; while negative reactions are observed for amidon, amygdalin, arbutin, D-adonitol, D-arabitol, D-cellobiose, D-lactose, D-lyxose, D-maltose, D-mannitol, D-melezitose, D-melibiose, D-raffinose, D-saccharose, D-sorbitol, D-tagatose, D-trehalose, D-turanose, dulcitol, erythritol, aesculin, gentiobiose, glycerol, glycogen, inositol, inulin, L-arabitol, L-fucose, L-rhamnose, L-sorbose, L-xylose, methyl- $\alpha$ -D-glucopyranoside, methyl- $\alpha$ -D-mannopyranoside, methyl- $\beta$ -D-xylopyranoside, *N*-acetylglucosamine, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, salicin and xylitol. The main fatty acids are C<sub>16:0</sub> and summed feature 8, comprising C<sub>18:1</sub>  $\omega$ 7c and/or C<sub>18:1</sub>  $\omega$ 6c. The main respiratory quinone is Q-9.

The type strain, CPA62<sup>T</sup>, was isolated from the surface sterilized aboveground tissues of the halophyte *Halimione portulacoides*. The G+C content of the type strain, obtained from genomic data, is 63.5 mol%.

### **Description of *Salinicola endophytica* sp. nov. CPA92<sup>T</sup>**

*Salinicola endophytica* (Gr. pref. *endo*, within; Gr. n. *phyton*, plant; L. fem. suff. *-ica*, adjectival suffix used with the sense of belonging to; N.L. fem. adj. *endophytica*, within plant, endophytic).

The CPA92<sup>T</sup> strain is a Gram-negative, positive for catalase and oxidase, facultative anaerobe, motile rod (1.50–3.75  $\mu$ m L  $\times$  0.73–1.00  $\mu$ m W). Colonies are pale yellow, opaque in the center and less so around the regular smooth edges, with 1.2–2 mm in diameter, after incubation in MA for 48 h at 30 °C. Growth is observed at salinities of 0 to 25 % NaCl (w/v) with a salinity optimum at 7.5 % NaCl (w/v); at temperatures ranging from 4 to 45 °C, with optimum growth observed at 30 °C; and from pH 4.0 to 10.0, where the optimum is observed in the 6.0–7.0 interval. Production of H<sub>2</sub>S is not observed. Cellulose, citrate, starch, Tween 20 and xylan are hydrolyzed, however, casein and pectin are not. DNase activity is weak. Phosphate solubilization, IAA production (29.7  $\mu$ g mL<sup>-1</sup>), siderophore production and ACC deaminase activity are observed. In the API 20NE strip, the strain is positive for urease, and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetylglucosamine, D-maltose, potassium gluconate, capric acid, malic acid and trisodium citrate; and negative for reduction of nitrates to nitrites, indole production, fermentation of D-glucose, arginine dihydrolase,  $\beta$ -glucosidase (aesculin), protease (gelatin),  $\beta$ -galactosidase (para-nitrophenyl- $\beta$ -D-galactopyranose), and assimilation of adipic acid and phenylacetic acid. In the API

ZYM strip, the strain is positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase,  $\alpha$ -glucosidase; and negative for lipase (C14), cysteine arylamidase, trypsin,  $\alpha$ -chymotrypsin,  $\alpha$ -galactosidase,  $\beta$ -galactosidase,  $\beta$ -glucuronidase,  $\beta$ -glucosidase, *N*-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -mannosidase and  $\alpha$ -fucosidase. In API 50CH, the strain utilizes D-arabinose, D-fructose, D-fucose, D-galactose, D-glucose, D-lyxose, D-mannitol, D-mannose, D-ribose, D-sorbitol, D-turanose, dulcitol, D-xylose, erythritol and L-arabinose. The strain weakly utilized D-cellobiose, D-lactose, D-melibiose, D-saccharose, D-tagatose, L-rhamnose and L-sorbose, and did not utilize amidon, amygdalin, arbutin, D-adonitol, D-arabitol, D-maltose, D-melezitose, D-raffinose, D-trehalose, aesculin, gentiobiose, glycerol, glycogen, inositol, inulin, L-arabitol, L-fucose, L-xylose, methyl- $\alpha$ -D-glucopyranoside, methyl- $\alpha$ -D-mannopyranoside, methyl- $\beta$ -D-xylopyranoside, *N*-acetylglucosamine, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, salicin and xylitol. The main fatty acids are C<sub>16:0</sub> and summed feature 8, comprising C<sub>18:1</sub>  $\omega$ 7c and/or C<sub>18:1</sub>  $\omega$ 6c. The main respiratory quinone is Q-9.

The type strain, CPA92<sup>T</sup>, was isolated from the surface sterilized aboveground tissues of the halophyte *Halimione portulacoides*. The G+C content of the type strain, obtained from genomic data, is 65.8 mol%.

### **Description of *Salinicola halophytica* sp. nov. CR45<sup>T</sup>**

*Salinicola halophytica* (N.L. fem. adj. *halophytica*, halophytic, pertaining to the original isolation from halophyte tissues).

The CR45<sup>T</sup> strain is a Gram-negative, positive for catalase and oxidase, facultative anaerobe, motile rod (1.76–2.59  $\mu$ m L  $\times$  0.70–0.90  $\mu$ m W). Colonies are pale yellowish pink, opaque in the center and less so around the regular smooth edges, with 1.6–2 mm in diameter, after incubation in MA for 48 h at 30 °C. Salt is necessary for growth, and the NaCl tolerance ranges from 2.5 to 20 % NaCl (w/v) with an optimum salinity in the interval of 7.5 to 10 % NaCl (w/v); at temperatures ranging from 4 to 40 °C, with optimum growth observed at 25 °C; and from pH 5.0 to 10.0, where the optimum is observed at pH 6.0. Production of H<sub>2</sub>S is not observed. Pectin (at pH 7.0), Tween 20 and citrate are hydrolyzed, but casein, cellulose, starch and xylan are not. No growth is observed in the medium to test pectin hydrolysis at pH 5.0. DNase activity is weak. Phosphate solubilization, IAA production (27.3  $\mu$ g mL<sup>-1</sup>), siderophore production and ACC deaminase activity are observed. In the API 20NE

strip, the strain is positive for the assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol and potassium gluconate; and negative for reduction of nitrates to nitrites, indole production, fermentation of D-glucose, arginine dihydrolase, urease,  $\beta$ -glucosidase (aesculin), protease (gelatin),  $\beta$ -galactosidase (para-nitrophenyl- $\beta$ -D-galactopyranose), assimilation of *N*-acetylglucosamine, D-maltose, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid. In the API ZYM strip, the strain is positive for alkaline phosphatase, leucine arylamidase, valine arylamidase and acid phosphatase; and weakly positive for naphthol-AS-BI-phosphohydrolase. The strain is negative for esterase (C4), esterase lipase (C8), lipase (C14), cysteine arylamidase, trypsin,  $\alpha$ -chymotrypsin,  $\alpha$ -galactosidase,  $\beta$ -galactosidase,  $\beta$ -glucuronidase,  $\alpha$ -glucosidase,  $\beta$ -glucosidase, *N*-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -mannosidase and  $\alpha$ -fucosidase. In the API 50CH strip, the strain utilizes D-cellobiose, D-fructose, D-fucose, D-galactose, D-glucose, D-mannose, D-ribose, D-xylose and L-arabinose. The strain weakly utilizes D-lactose, D-maltose and aesculin; and does not utilize amidon, amygdalin, arbutin, D-adonitol, D-arabinose, D-arabitol, D-lyxose, D-mannitol, D-melezitose, D-melibiose, D-raffinose, D-saccharose, D-sorbitol, D-tagatose, D-trehalose, D-turanose, dulcitol, erythritol, gentiobiose, glycerol, glycogen, inositol, inulin, L-arabitol, L-fucose, L-rhamnose, L-sorbose, L-xylose, methyl- $\alpha$ -D-glucopyranoside, methyl- $\alpha$ -D-mannopyranoside, methyl- $\beta$ -D-xylopyranoside, *N*-acetylglucosamine, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, salicin and xylitol. The main fatty acids are C<sub>16:0</sub>, summed feature 8, comprising C<sub>18:1</sub>  $\omega$ 7c and/or C<sub>18:1</sub>  $\omega$ 6c, and cyclo-C<sub>19:0</sub>  $\omega$ 8c. The main respiratory quinone is Q-9.

The type strain, CR45<sup>T</sup>, was isolated from the surface sterilized belowground tissues of the halophyte *Halimione portulacoides*. The G+C content of the type strain, obtained from genomic data, is 62.2 mol%.

### **Description of *Salinicola lusitana* sp. nov. CR50<sup>T</sup>**

*Salinicola lusitana* (L. fem. adj. *lusitana*, pertaining to *Lusitania*, the Roman name of Portugal, where it was first isolated).

The CR50<sup>T</sup> strain is a Gram-negative, positive for catalase and oxidase, facultative anaerobe, motile rod (1.50–3.05  $\mu$ m L  $\times$  0.77–1.06  $\mu$ m W). Colonies are pale yellow, opaque in the center and less so around the smooth edges, with 1–2.1 mm in diameter, after incubation in MA for 48 h at 30 °C. NaCl tolerance ranges from 0 to 25 % NaCl (w/v) with a salinity optimum in the interval of 5 to 7.5

% NaCl (w/v); at temperatures ranging from 15 to 45 °C, with optimum growth observed at 30 °C; and from pH 4.0 to 10.0, where the optimum is observed at pH 6.0. Production of H<sub>2</sub>S is not observed. Cellulose, DNA, pectin (at pH 7.0), starch, Tween 20 and citrate are hydrolyzed, but casein, pectin (pH 5.0) and xylan are not. Phosphate solubilization, IAA production (45.7 µg mL<sup>-1</sup>), siderophore production and ACC deaminase activity are observed. In the API 20NE strip, the strain is positive for assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetylglucosamine, D-maltose, potassium gluconate, capric acid, malic acid and trisodium citrate; and negative for reduction of nitrates to nitrites, indole production, fermentation of D-glucose, arginine dihydrolase, urease, β-glucosidase (aesculin), protease (gelatin), β-galactosidase (para-nitrophenyl-β-D-galactopyranose), assimilation of adipic acid and phenylacetic acid. In the API ZYM strip, the strain is positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, and α-glucosidase. The strain is negative for lipase (C14), cysteine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase. In the API 50CH strip, the strain utilizes D-fucose, D-galactose, D-glucose, D-ribose, D-xylose and L-arabinose. The strain weakly utilizes D-arabinose, D-fructose, D-mannose and erythritol; and does not utilize amidon, amygdalin, arbutin, D-adonitol, D-arabitol, D-cellobiose, D-lactose, D-lyxose, D-maltose, D-mannitol, D-melezitose, D-melibiose, D-raffinose, D-saccharose, D-sorbitol, D-tagatose, D-trehalose, D-turanose, dulcitol, aesculin, gentiobiose, glycerol, glycogen, inositol, inulin, L-arabitol, L-fucose, L-rhamnose, L-sorbose, L-xylose, methyl-α-D-glucopyranoside, methyl-α-D-mannopyranoside, methyl-β-D-xylopyranoside, *N*-acetylglucosamine, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, salicin and xylitol. The main fatty acids are C<sub>16:0</sub>, summed feature 8, comprising C<sub>18:1</sub> ω7c and/or C<sub>18:1</sub> ω6c, and cyclo-C<sub>19:0</sub> ω8c. The main respiratory quinone is Q-9.

The type strain, CR50<sup>T</sup>, was isolated from the surface sterilized belowground tissues of the halophyte *Halimione portulacoides*. The G+C content of the type strain, obtained from genomic data, is 64.3 mol%.

## Acknowledgements

This work was supported by European Funds (FEDER) through COMPETE and by National Funds through the Portuguese Foundation for Science and Technology (FCT) within project PhytoMarsh (PTDC/AAC-AMB/118873/2010 – FCOMP-01-0124-FEDER-019328). FCT is acknowledged for financing to CESAM (UID/AMB/50017/2013), Artur Alves (FCT Investigator Programme – IF/00835/2013), Isabel Henriques (FCT Investigator Programme – IF/00492/2013), Cátia Fidalgo (PhD grant – SFRH/BD/85423/2012), and Diogo Neves Proença (postdoctoral fellowship – SFRH/BPD/100721/2014).

## References

- Aguilera, M., Cabrera, A., Incerti, C., Fuentes, S., Russell, N. J., Ramos-Cormenzana, A. & Monteoliva-Sánchez, M. (2007). *Chromohalobacter salarius* sp. nov., a moderately halophilic bacterium isolated from a solar saltern in Cabo de Gata, Almería, southern Spain. *Int J Syst Evol Microbiol* **57**, 1238-1242.
- Anan'ina, L. N., Plotnikova, E. G., Gavrish, E. Y., Demakov, V. A. & Evtushenko, L. I. (2007). *Salinicola socius* gen. nov., sp. nov., a Moderately Halophilic Bacterium from a Naphthalene-Utilizing Microbial Association. *Microbiology* **76**(3), 324-330.
- Andrews, S. (2010). FastQC: A quality control tool for high-throughput sequence data. Available online at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Arahal, D. R., Ludwig, W., Schleifer, K. H. & Ventosa, A. (2002a). Phylogeny of the family *Halomonadaceae* based on 23S and 16S rDNA sequence analyses. *Intern Int J Syst Evol Microbiol* **52**, 241-249.
- Arahal, D. R., Castillo, A. M., Ludwig, W., Schleifer, K. H. & Ventosa, A. (2002b). Proposal of *Cobetia marina* gen. nov., comb. nov., within the Family *Halomonadaceae*, to include the species *Halomonas marina*. *Syst Appl Microbiol* **25**, 207-211.
- Arahal, D. R., Vreeland, R. H., Litchfield, C. D., Mormile, M. R., Tindall, B. J., Oren, A., Bejar, V., Quesada, E. & Ventosa, A. (2007). Recommended minimal standards for describing new taxa of the family *Halomonadaceae*. *Int J Syst Evol Microbiol* **57**, 2436-2446.
- Aziz, R.K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**, 75.
- Cao, L., Yan, Q., Ni, H., Hu, G., Hong, Q. & Li, S. (2013). *Salinicola zeshunii* sp. nov., a moderately halophilic bacterium isolated from soil of a chicken farm. *Curr Microbiology* **66**, 192-196.

- de la Haba, R. R., Arahal, D. R., Márquez, M. C. & Ventosa, A. (2010a).** Phylogenetic relationships within the family *Halomonadaceae* based on comparative 23S and 16S rRNA gene sequence analysis. *Int J Syst Evol Microbiol* **60**, 737-748.
- de la Haba, R. R., Sánchez-Porro, C., Márquez, M. C. & Ventosa, A. (2010b).** Taxonomic study of the genus *Salinicola*: transfer of *Halomonas salaria* and *Chromohalobacter salarius* to the genus *Salinicola* as *Salinicola salarius* comb. nov. and *Salinicola halophilus* nom. nov., respectively. *Int J Syst Evol Microbiol* **60**, 963-971.
- de la Haba, R. R., Márquez, M. C., Papke, R. T. & Ventosa, A. (2012).** Multilocus sequence analysis of the family *Halomonadaceae*. *Int J Syst Evol Microbiol* **62**, 520-538.
- Dobson, S. J. & Franzmann, P. D. (1996).** Unification of the genera *Deleya* (Baumann et al. 1983), *Halomonas* (Vreeland et al. 1980), and *Halovibrio* (Fendrich 1988) and the species *Paracoccus halodenitrificans* (Robinson and Gibbons 1952) into a single genus, *Halomonas*, and placement of the genus *Zymobacter* in the Family *Halomonadaceae*. *Int J Syst Bacteriol* **46**(2), 550-558.
- Felsenstein, J. (1981).** Evolutionary trees from DNA sequences: a maximum-likelihood approach. *J Mol Evol* **17**, 368-376.
- Fidalgo, C., Henriques, I., Rocha, J., Tacão, M. & Alves, A. (2016).** Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization, and influence of metal(loid) contamination. *Environ Sci Pollut Res* **23**(10), 10200-10214.
- Figueras, M. J., Beaz-Hidalgo, R., Hossain, M. J. & Liles, M. R. (2014).** Taxonomic affiliation of new genomes should be verified using average nucleotide identity and multilocus phylogenetic analysis. *Genome Announc* **2**, e00927-14.
- Fitch, W. M. (1971).** Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* **20**, 406-416.
- Franzmann, P. D., Wehmeyer, U. & Stackebrandt, E. (1988).** *Halomonadaceae* fam. nov., a new family of the class Proteobacteria to accommodate the genera *Halomonas* and *Deleya*. *Syst Appl Microbiol* **11**, 16-19.
- Gam, Z. B. A., Abdelkafi, S., Casalot, L., Tholozan, J. L., Oueslati, R. & Labat, M. (2007).** *Modicisalibacter tunisiensis* gen. nov., sp. nov., an aerobic, moderately halophilic bacterium isolated from an oilfield-water injection sample, and emended description of the family *Halomonadaceae* Franzmann et al. 1989 emend Dobson and Franzmann 1996 emend. Ntougias et al. 2007. *Int J Syst Evol Microbiol* **57**, 2307-2313.
- Hall, T. A. (1999).** BioEdit: a user-friendly biological sequence alignment edit and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95-98.
- Huo, Y.-Y., Meng, F.-X., Xu, L., Wang, C.-S., Xu & X.-W. (2013).** *Salinicola peritrichatus* sp. nov., isolated from deep-sea sediment. *Antonie van Leeuwenhoek* **104**, 55-62.



- Kim, K. K., Jin, L., Yang, H. C. & Lee, S.-T. (2007).** *Halomonas gomseomensis* sp. nov., *Halomonas janggokensis* sp. nov., *Halomonas salaria* sp. nov. and *Halomonas denitrificans* sp. nov., moderately halophilic bacteria isolated from saline water. *Int J Syst Evol Microbiol* **57**, 675-681.
- Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H., Yi, H., Won, S. & Chun, J. (2012).** Introducing EzTaxon: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* **62**, 716–721.
- Kim, M., Park, S.-C., Baek, I. & Chun, J. (2015).** Large-scale evaluation of experimentally determined DNA G+C contents with whole genome sequences of prokaryotes. *Syst Appl Microbiol* **38**, 79-83.
- Kimura, M. (1980).** A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Lagesen, K., Hallin, P., Rodland, E. A., Staerfeldt, H.-H., Rognes, T. & Ussery, D. W. (2007).** RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* **35**, 3100–3108.
- León, M. J., Sánchez-Porro, C., de la Haba, R. R., Llamas, I. & Ventosa, A. (2014).** *Larsenia salina* gen. nov., sp. nov., a new member of the family *Halomonadaceae* based on multilocus sequence analysis. *Syst Appl Microbiol* **37**, 480-487.
- Lepcha, R. T., Poddar, A., Schumann, P. & Das, S. K. (2015).** Comparative 16S rRNA signatures and multilocus sequence analysis for the genus *Salinicola* and description of *Salinicola acroporae* sp. nov., isolated from coral *Acropora digitifera*. *Antonie van Leeuwenhoek* **108**, 59-73.
- Mata, J. A., Martínez-Cánovas, J., Quesada, E. & Béjar, V. (2002).** A detailed phenotypic characterisation of the type strains of *Halomonas* species. *Syst Appl Microbiol* **25**, 360-375.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013).** Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **41**, (Web Server issue):W597-600 doi:10.1093/nar/gkt376.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P. & Göker, M. (2013).** Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**, 60.
- Meier-Kolthoff, J. P., Klenk, H.-P. & Göker, M. (2014).** Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* **64**, 352-356.
- Mellado, E., Moore, E. R. B., Nieto, J. J. & Ventosa, A. (1995).** Phylogenetic inferences and taxonomic consequences of 16S ribosomal DNA sequence comparison of *Chromohalobacter marismortui*, *Volcaniella eurihalina*, and *Deleya salina* and reclassification of *V. eurihalina* as *Halomonas eurihalina* comb. nov.. *Int J Syst Bacteriol* **45**(4), 712-716.

**Mesbah, M., Premachandran, U. & Whitman, W. B. (1989).** Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* **39**, 159–167.

**NCBI Resource Coordinators (2016).** Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**, D7–D19. <http://doi.org/10.1093/nar/gkv1290>

**Ntougias, S., Zervakis, G. I. & Fasseas, C. (2007).** *Halotalea alkalilenta* gen. nov., sp. nov., a novel osmotolerant and alkalitolerant bacterium from alkaline olive mill wastes, and emended description of the family *Halomonadaceae* Franzmann et al. 1989, emend. Dobson and Franzmann 1996. *Int J Syst Evol Microbiol* **57**, 1975–1983.

**Okamoto, T., Taguchi, H., Nakamura, K., Ikenaga, H., Kuraishi, H. & Yamasato, K. (1993).** *Zymobacter palmae* gen. nov., sp. nov., a new ethanol-fermenting peritrichous bacterium isolated from palm sap. *Arch Microbiol* **160**, 333–337.

**Oren, A. & Ventosa, A. (2013).** Subcommittee on the taxonomy of *Halobacteriaceae* and Subcommittee on the taxonomy of *Halomonadaceae*. *Int J Syst Evol Microbiol* **63**, 3540–3544.

**Oren, A. & Ventosa, A. (2016).** International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of *Halobacteriaceae* and subcommittee on the taxonomy of *Halomonadaceae*. Minutes of the joint open meeting, 23 May 2016, San Juan, Puerto Rico. *Int J Syst Evol Microbiol* **66**, 4291–4295.

**Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F. & Stevens, R. (2014).** The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acid Res* **42**, D206–214.

**Owen, R. J. & Pitcher, D. (1985).** Current methods for estimating DNA base composition and levels of DNA–DNA hybridization. In *Chemical Methods in Bacterial Systematics*, pp. 67–93. Edited by M. Goodfellow & D. E. Minnikin. London: Academic Press.

**Proença, D. N., Nobre, M. F. & Morais, P. V. (2014).** *Chitinophaga costaii* sp. nov., an endophyte of *Pinus pinaster*, and emended description of *Chitinophaga niabensis*. *Int J Syst Evol Microbiol* **64**, 1237–1243.

**Raju, K., Sekar, J. & Ramalingam, P. V. (2016).** *Salinicola rhizosphaerae* sp. nov., isolated from the rhizosphere of the mangrove *Avicennia marina* L. *Int J Syst Evol Microbiol* **66**, 1074–1079.

**Richter, M. & Rosselló-Móra, R. (2009).** Shifting the genomic gold standard for the prokaryotic species definition. *PNAS* **106**, 19126–19131.

**Richter, M., Rosselló-Móra, R., Glöckner, F. O. & Peplies, J. (2016).** JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **32**, 929–931.

**Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

**Sánchez-Porro, C., de la Haba, R. R., Soto-Ramírez, N., Márquez, M. C., Montalvo-Rodríguez, R. & Ventosa, A. (2009).** Description of *Kushneria aurantia* gen. nov., sp. nov., a novel member of the family *Halomonadaceae*, and a proposal for reclassification of *Halomonas marisflavi* as *Kushneria marisflavi* comb. nov., of *Halomonas indalinina* as *Kushneria indalinina* comb. nov. and of *Halomonas avicenniae* as *Kushneria avicenniae* comb. nov.. *Int J Syst Evol Microbiol* **59**, 397-405.

**Stackebrandt, E. & Goebel, B. M. (1994).** Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* **44**, 846-849.

**Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013).** MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.

**Ventosa, A., Quesada, E., Rodriguez-Valera, F., Ruiz-Berraquero, F. & Ramos-Cormenzana, A. (1982).** Numerical taxonomy of moderately halophilic Gram-negative rods. *J Gen Microbiol* **128**, 1959-1968.

**Villesen, P. (2007).** FaBox: an online toolbox for fasta sequences. *Mol Ecol Notes* **7**, 965–968.

**Vreeland, R. H. & Ventosa, A. (2003).** International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of the *Halomonadaceae*. *Int J Syst Evol Microbiol* **53**, 921-922.

**Wang, Y., Tang, S.-K., Lou, K., Lee, J.-C., Jeon, C. O., Xu, L.-H., Kim, C.-J. & Li, W.-J. (2009).** *Aidingimonas halophila* gen. nov., sp. nov., a moderately halophilic bacterium isolated from a salt lake. *Int J Syst Evol Microbiol* **59**, 3088-3094.

**Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. & Trüper, H. G. (1987).** Report of the Ad Hoc Committee on reconciliation of approaches to bacterial systematics. *Int J Syst Evol Microbiol* **37**, 463-464.



## Chapter 5

---

Illumina-based analysis of endophytic bacteria from *Halimione portulacoides*

## **Contents**

Abstract

Background

Motivation

Methods

Results and discussion

Conclusions

References

Annex I

Annex II

## Abstract

Fifteen healthy specimens of *Halimione portulacoides* were collected from a salt marsh in Ria de Aveiro. Surface-sterilized aboveground (AG) and belowground (BG) tissues of the halophyte were analyzed for their diversity of endophytic bacteria using a culture-independent high-throughput 16S rRNA gene sequencing method. The interference of host DNA in the analysis was decreased by including PNA blockers in the 16S rRNA gene amplification reaction. The methodology used allowed for a high sequencing coverage for all samples (median 97.5 %) and revealed a high operational taxonomic units (OTUs) richness (median of 200 OTUs per sample). Results showed significant differences according to sampling tissue, where BG tissues presented higher OTU richness and diversity. Community structure was also found to differ according to sampling tissue, and this difference may be associated with a large number of rare OTUs. Overall, the main phyla observed in the endosphere of the halophyte were Proteobacteria (62.5 % OTUs), Bacteroidetes (10.7 %), Planctomycetes (8.8 %), Actinobacteria (5.2 %) and Firmicutes (2.4 %). Core endophytome analyses revealed that the AG tissues mainly comprised members of the family *Oceanospirillaceae*, and also included the families *Flammeovirgaceae*, *Enterobacteriaceae* and *Flavobacteriaceae*, while the BG endophytome was dominated by *Enterobacteriaceae* and *Kiloniellaceae*. This work explored the diversity of the bacterial endophytome of *H. portulacoides* in depth and confirmed it as a bacterial hotspot. Putative ecological functions of the dominant taxa revealed a community that presented several plant growth promotion traits.

## Background

Plant-associated microbes are ecologically relevant due to their roles in various processes including nutrient acquisition and cycling, protection against pathogens, and production of traits that allow the plant to thrive (Hardoim et al., 2015; Santoyo et al., 2016). In order to understand the potential of a plant host, it should be seen as a holobiont, thus requiring an extensive knowledge of the structure and diversity of the complex host-associated microbial community (reviewed in Vandenkoornhuyse et al., 2015). Culture-based methodologies have been thoroughly used to characterize such communities, however, this approach relies on the ability of the microbes to grow in laboratory conditions. This conditioning, whether in terms of nutrient availability, temperature, pressure, or other conditions, usually results in a collection of microbes that represents a minor fraction of the community (Amann et al., 1995; Tanaka et al., 2014).

Recent advances in amplicon-based high-throughput sequencing and its broad application have allowed for a more comprehensive assessment of plant-associated bacterial communities. The focus of these studies has been largely aimed at the microbes present in the rhizospheric portion, since this type of environment is very abundant in ecologically relevant microbes (Philippot et al., 2013). The endosphere, albeit also rich in relevant microbes (Hardoim et al., 2015), has been less explored. The reason for this may be attributed to the impairment caused by the presence of plant organelle DNA in the samples used in endophytic bacterial diversity assessments. In healthy plant specimens, these samples present a low ratio of bacterial:host cells, and the latter may contain up to tens of thousands of organelle genomes in each cell, depending on the plant species (Bendich, 1987). Since organelle genomes contain DNA sequences similar to the widely used 16S rRNA gene for bacterial identification, the molecular-based techniques often result in an output rich in organelle reads, and less so in bacterial reads.



## Motivation

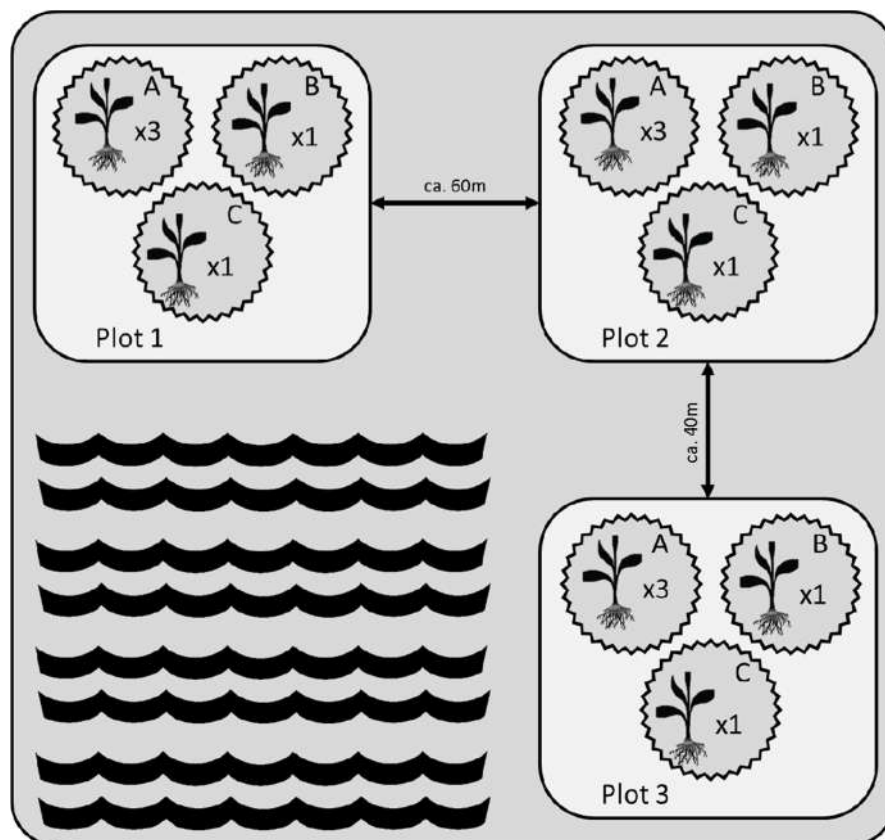
Our assessment of culturable endophytic bacteria from *Halimione portulacoides* yielded 665 isolates (Chapter 3), and it is considered that the culturable fraction shows a small fraction of the diversity within a microbial community (Amann et al., 1995). Since the main purpose of the present work was to have an understanding of the bacterial diversity in the halophyte *H. portulacoides*, it was our aim to explore this diversity using a culture-independent, molecular-based, high-throughput technique. Since the diversity of culturable bacteria exhibited significant differences in aboveground and belowground tissues, here we also focused on the exploration of these compartments separately.

Our aim was, then, to achieve a deeper understanding of the bacterial diversity present in the surface-sterilized AG and BG tissues of *H. portulacoides* collected in Ria de Aveiro. This will allow us to determine the core bacterial endophytome associated with the halophyte.

## Methods

### Study area and sampling

Sampling was performed in a salt marsh in the estuary Ria de Aveiro (40° 38' 05.36" N, 8° 39' 38.87" W; site C from Chapter 3) during low tide in December 2015 (average temperature 11 °C). Fifteen healthy specimens of the halophyte *H. portulacoides* were collected in monospecific stands, in the design exposed in Figure 5.1. Three roughly equidistant plots were chosen for sampling. From each of the three plots, five whole plants were extracted from three monospecific stands of *H. portulacoides*: three plants from stand A, one plant from stand B and one plant from stand C. This differential sampling was performed to understand if the community was affected by the sampling stand. Samples were transported to the laboratory and promptly processed.



**Figure 5.1** Schematics of the sampling strategy to obtain fifteen specimens of *Halimione portulacoides* from Ria de Aveiro. Three roughly equidistant plots were chosen (Plot 1, 2 and 3), and in each plot three monospecific stands of *H. portulacoides* were sampled: in the first stand (stand A) three specimens of the halophyte were collected, and one specimen from stand B and another from stand C were collected.

### Sample processing, DNA extraction, 16S rRNA gene amplification and amplicon sequencing

In the laboratory, the fifteen plants were washed with running tap water and the AG tissues were separated from BG tissues, resulting in 30 individual samples. A detailed description of the 30 samples is given in Table 5.1.

**Table 5.1** Description of the samples used in the study.

Sample	Plot	Sample description		Tissue
		Stand	Specimen	
1	1	A	1	AG
2	1	A	1	BG
3	1	A	2	AG
4	1	A	2	BG
5	1	A	3	AG
6	1	A	3	BG
7	1	B	4	AG
8	1	B	4	BG
9	1	C	5	AG
10	1	C	5	BG
11	2	A	1	AG
12	2	A	1	BG
13	2	A	2	AG
14	2	A	2	BG
15	2	A	3	AG
16	2	A	3	BG
17	2	B	4	AG
18	2	B	4	BG
19	2	C	5	AG
20	3	C	5	BG
21	3	A	1	AG
22	3	A	1	BG
23	3	A	2	AG
24	3	A	2	BG
25	3	A	3	AG
26	3	A	3	BG
27	3	B	4	AG
28	3	B	4	BG
29	3	C	5	AG
30	3	C	5	BG

AG, aboveground tissues; BG, belowground tissues.

Surface sterilization was performed as described before (Chapter 3) and aliquots (100  $\mu$ L) of water from the last wash were used as inoculum in Tryptic Soy Agar (TSA, Merck, Germany) medium. The absence of growth after 7 days incubating at 30 °C confirmed the efficacy of the sterilization procedure.

Portions of the samples were macerated using a micro-pestle in a 1.5 mL Eppendorf tube. DNA extraction was conducted using the MoBio PowerPlant DNA Isolation kit (USA) according to the manufacturer's instructions with the exception that the DNA was eluted in PCR-grade water (Sigma, USA) in the final step. DNA concentration was determined spectrophotometrically using NanoDrop ND-1000 (Thermo Scientific, USA). PCR amplification of the 16S rRNA gene was performed using 25 ng of template DNA, 0.2  $\mu$ M of primer 341F (5'-CCTACGGGNGGCWGCAG-3'; Muyzer et al., 1993) and 805R (5'-GACTACHVGGGTATCTAATCC-3'; Skirnisdottir et al., 2000), 2  $\mu$ M of mPNA and pPNA (Lundberg et al., 2013; PNAbio, USA), and 2X KAPA HiFi Ready Mix (KAPA Biosystems, USA). Prior to using the mPNA and pPNA in the reaction, these reagents were allowed to solubilize at 55 °C for 30 min. Conditions for amplification were as follows: one cycle at 95 °C (3 min), followed by 30 cycles at 98 °C (20 sec), 75 °C (10 sec), 55 °C (30 sec) and 72 °C (30 sec), with a final extension at 72 °C (5 min). The PCR amplification was also conducted on a randomly chosen sample (sample 11) without addition of the mPNA and pPNA reagents, in order to understand the effect of the PNA blockers in the Illumina MiSeq sequencing results. This sample is regarded from now on as sample -PNA, and the 30 samples for which the PCR reactions were performed with addition of PNA blockers are regarded as +PNA.

At the sequencing company (Biocant, Portugal), the samples were subjected to an electrophoresis on agarose gel [1 % (w/v)] and the 490 bp amplicons were purified using AMPure XP beads (Agencourt, USA) according to manufacturer's instructions. A second PCR reaction added indexes and sequencing adaptors to both ends of the amplified target region using 2X KAPA HiFi HotStart Ready Mix, 5  $\mu$ L of each index (i7 and i5) (Nextera XT Index Kit, Illumina, USA) and 5  $\mu$ L of the first PCR product in a total volume of 50  $\mu$ L. The PCR conditions were as follows: 95 °C (3 min), eight cycles of 95 °C (30 sec), 55 °C (30 sec) and 72 °C (30 sec), and a final extension at 72 °C (5 min). Electrophoresis of the PCR products was undertaken on an agarose gel [1 % (w/v)] and the 550 bp amplified fragments were purified using AMPure XP beads (Agencourt, USA) according to manufacturer's instructions. The amplicons were quantified by fluorimetry with PicoGreen dsDNA quantitation kit (Invitrogen, USA), pooled at equimolar concentrations and pair-end sequenced using an Illumina MiSeq V3 platform, according to manufacturer's instructions (Illumina, USA).

## Data processing and analysis

At the sequencing company, the sequenced reads were demultiplexed automatically by the Illumina Miseq sequencer using the CASAVA package (Illumina, USA) and quality-filtered with PRINSEQ software (Schmieder and Edwards, 2011) using the following parameters: (i) bases with average quality lower than Q25 in a window of 5 bases were trimmed, and (ii) reads with less than 150 bases were discarded. The forward and reverse reads were merged by overlapping paired-end reads using the AdapterRemoval v2.1.5 (Schubert et al., 2016) software with default parameters.

The QIIME pipeline (Caporaso et al., 2010) was used to further process the reads. For the +PNA samples, the 30 separate sequence files were merged into a single fasta file using the script `split_libraries_fastq.py`. The files for -PNA and +PNA samples were then searched for chimeras with the `usearch61` method using the script `identify_chimeric_seqs.py` (Edgar, 2010) and the chimeric sequences were filtered out using the script `filter_fasta.py`. OTU tables with taxonomic affiliation were then created using the script `pick_open_reference_otus.py` (Rideout et al., 2014) and the Greengenes database (DeSantis et al., 2006). The OTU table that resulted from the -PNA sample was not further studied.

The OTU table that resulted from the +PNA samples was then imported to (function `read.csv`; package `utils`) and analyzed with R software (R Core Team, 2016). Alpha-diversity was explored using several metrics included in the `vegan` package (Oksanen et al., 2016): Shannon's diversity ( $H$ ; function `diversity`), richness (function `specnumber`), Pielou's evenness ( $J$ ; obtained by calculating  $H/\log_{10}(\text{richness})$ ), Simpson's diversity ( $1-D$ ; function `diversity`, index `simpson`), and inverse Simpson's diversity ( $1/D$ ; function `diversity`, index `invsimpson`). Additionally, the sequencing coverage of all samples was estimated using Chao's coverage estimator, which takes into account the presence of singletons and doubletons (function `coverage`, package `entropart`; Marcon & Hérault, 2015). Each alpha-diversity metric was then statistically analyzed using functions from the package `stats` (R Core Team, 2016). First, the normality of the distribution of the values from all samples was tested using the Shapiro normality test (function `shapiro.test`). When normality was observed, (i) an ANOVA (function `aov`) was performed to assess statistical differences in this metric among the different plots; (ii) a t-Student test (function `t.test`) allowed determination of statistical difference among the original plant tissues. If normality was not observed, (i) plot-related differences were assessed by performing a Kruskal-Wallis rank sum test (function `kruskal.test`), and (ii) tissue-related differences were tested with the Wilcoxon rank sum test (function `wilcox.test`).

Rarefaction curves were obtained using the function *rarecurve* (package *vegan*) and allowed assessment of the sequencing depth for each sample.

Transformation of the OTU table was performed with the functions *sqrt* and *log10*, both from the package *base* (R Core Team, 2016). Resemblance matrices of these tables were built using the *vegdist* function from the package *vegan* with the method Bray-Curtis. A permutational multivariate analysis of variance using distance matrices (PERMANOVA) was performed using the function *adonis* from the package *vegan* to assess the effect of the factors plot and tissue on the OTU tables. For instances where differences are observed among the different plots, the function *pairwise.adonis* was used to determine the plot pairs that were effectively different (Annex I). Ordination methods were performed with the functions *metaMDS* and *cmdscale*, both from the package *vegan*, to perform non-metric multidimensional scaling (NMDS) and principal coordinates analysis (PCoA), respectively.

The MetaCoMET platform was used to perform analyses on the occurrence of OTUs among tissues and sampling plots (Venn diagrams), based on the same OTU data that was analyzed with R (Wang et al., 2016b). MetaCoMET was also used to further investigated OTUs that were shared by and exclusive to halophyte tissues or sampling plots, using as input: all samples, only AG samples, or only BG samples. A Krona pie chart tool in this platform was used to visualize taxonomic hierarchical data in all samples.

## Results and discussion

### Samples and Illumina MiSeq sequencing

Fifteen specimens of *H. portulacoides* were sampled and analyzed in order to determine the core endophytic microbiome of the halophyte and tissue-specific taxa. For the +PNA samples, a total of 1761806 reads was obtained from Illumina MiSeq sequencing. Sequence filtering and processing steps reduced the number of reads: 1714836 reads after PRINSEQ quality filtering, 1584871 reads after merging paired-end sequences, 1550220 reads after demultiplexing, and 1538662 reads after filtering out chimeric sequences (read number per sample is detailed in Table 5.2). Overall, 87.33 % of the initial reads were kept after the filtering steps. For the -PNA sample, a total of 16229 reads was obtained after the sequence filtering and processing steps (data not shown).

### Host DNA contamination

One of the first steps in the study of endophytic communities involved total DNA extraction, which comprised DNA from different sources: plant nuclei DNA, plant organelle DNA (from chloroplasts and mitochondria) and bacterial DNA. The proportion of bacterial:host DNA depends on the plant species and other factors (Jiao et al., 2006). Due to the universal nature of the primers regularly used for amplification of the 16S rRNA gene, the most used gene in this type of studies, amplification of the 16S rRNA gene from chloroplast DNA and 18S rRNA gene from mitochondrial DNA is very common as a result of sequence similarity. Consequently, the OTU tables that researchers analyze after performing high-throughput sequencing are frequently populated with affiliations to organelle DNA.

**Table 5.2** Number of reads after each sequence data processing step.

Sample	Raw reads	Sequence data processing steps				
		Quality filtered (PRINSEQ)	Merged paired-end sequences	Demultiplexed sequences	Chimera filtered sequences	Filtered host DNA sequences
1	54629	53375 (97.7%)	49881 (91.31%)	49590 (90.78%)	49097 (89.87%)	28694 (52.53%)
2	39988	38896 (97.27%)	36322 (90.83%)	35865 (89.69%)	35664 (89.19%)	8299 (20.75%)
3	22506	21964 (97.59%)	20259 (90.02%)	20016 (88.94%)	19978 (88.77%)	228 (1.01%)
4	13814	13373 (96.81%)	12251 (88.69%)	12040 (87.16%)	11980 (86.72%)	2863 (20.73%)
5	28299	27595 (97.51%)	25220 (89.12%)	25085 (88.64%)	25036 (88.47%)	390 (1.38%)
6	49816	48955 (98.27%)	46089 (92.52%)	45806 (91.95%)	44922 (90.18%)	27130 (54.46%)
7	30788	30367 (98.63%)	28562 (92.77%)	28382 (92.19%)	28201 (91.6%)	16850 (54.73%)
8	29987	29179 (97.31%)	26753 (89.22%)	26391 (88.01%)	26346 (87.86%)	1060 (3.53%)
9	39493	37656 (95.35%)	34225 (86.66%)	33736 (85.42%)	33685 (85.29%)	479 (1.21%)
10	46458	45464 (97.86%)	41702 (89.76%)	41109 (88.49%)	41054 (88.37%)	1219 (2.62%)
11	39528	38704 (97.92%)	35769 (90.49%)	35376 (89.5%)	35318 (89.35%)	3905 (9.88%)
12	34173	33144 (96.99%)	30192 (88.35%)	29601 (86.62%)	29451 (86.18%)	4716 (13.8%)
13	48683	46935 (96.41%)	43238 (88.82%)	42647 (87.6%)	42571 (87.45%)	3811 (7.83%)
14	44305	43482 (98.14%)	39709 (89.63%)	39377 (88.88%)	39294 (88.69%)	636 (1.44%)
15	43877	42846 (97.65%)	39138 (89.2%)	39061 (89.02%)	38996 (88.88%)	175 (0.4%)
16	44724	43636 (97.57%)	39862 (89.13%)	39638 (88.63%)	39570 (88.48%)	1347 (3.01%)
17	80844	78655 (97.29%)	72833 (90.09%)	72481 (89.66%)	72364 (89.51%)	1028 (1.27%)
18	80249	76833 (95.74%)	70941 (88.4%)	65641 (81.8%)	65546 (81.68%)	1233 (1.54%)
19	65215	62593 (95.98%)	57099 (87.56%)	53363 (81.83%)	53237 (81.63%)	3318 (5.09%)
20	81085	79264 (97.75%)	70355 (86.77%)	69645 (85.89%)	68201 (84.11%)	59450 (73.32%)
21	87503	85083 (97.23%)	78956 (90.23%)	76482 (87.41%)	76371 (87.28%)	1153 (1.32%)
22	99863	97713 (97.85%)	91122 (91.25%)	89524 (89.65%)	88401 (88.52%)	41676 (41.73%)
23	92872	90434 (97.37%)	84727 (91.23%)	83050 (89.42%)	81901 (88.19%)	53385 (57.48%)
24	101649	99639 (98.02%)	93813 (92.29%)	92491 (90.99%)	89420 (87.97%)	66309 (65.23%)
25	73983	70583 (95.4%)	65026 (87.89%)	64743 (87.51%)	64570 (87.28%)	787 (1.06%)
26	92632	90466 (97.66%)	84232 (90.93%)	79891 (86.25%)	79176 (85.47%)	35521 (38.35%)
27	69061	67566 (97.84%)	61914 (89.65%)	61457 (88.99%)	61338 (88.82%)	448 (0.65%)
28	47464	45321 (95.48%)	41725 (87.91%)	40401 (85.12%)	40262 (84.83%)	2453 (5.17%)
29	89024	87567 (98.36%)	81830 (91.92%)	79723 (89.55%)	79388 (89.18%)	14060 (15.79%)
30	89294	87548 (98.04%)	81126 (90.85%)	77608 (86.91%)	77324 (86.59%)	10561 (11.83%)
Total	1761806	1714836 (97.33%)	1584871 (89.96%)	1550220 (87.99%)	1538662 (87.33%)	393184 (22.32%)



Common methods used to avoid host DNA contamination in culture-independent studies of endophytic bacteria communities have the main goal of shifting the bacterial:host DNA ratio in a manner that favors the amplification of bacterial template. These methods can intervene in different stages of the common steps in this type of study:

(i) During DNA extraction, where specific protocols are used so that bacterial and host DNA are separated (Jiao et al., 2006). Briefly, this method aims to differentially select bacterial cells for DNA extraction by physically separating the bacterial fraction from the host tissues by using a sucrose gradient and differential centrifugations. Prior to the centrifugations, an enzymatic digestion is performed on the plant tissues, so as not to disrupt plant cells, which would enhance host DNA contamination from chloroplasts and mitochondria. Jiao et al. (2006) showed that their method of enzyme treatment of plant tissues followed by differential centrifugation effectively increases the ratio of bacterial:host DNA. Nevertheless, this method requires a large amount of plant tissue, which is not always possible to obtain, depending on the plant species or tissue at study. Additionally, enzymatic incubation steps prior to the differential centrifugation may allow for bacterial proliferation, thus altering their initial structure (Wang et al., 2008). Furthermore, it follows reason that the centrifugation steps may obscure some bacterial diversity and alter the community structure by inadvertently discarding bacterial cells that remained attached to plant tissue debris.

(ii) After DNA extraction and prior to PCR amplification, using restriction enzymes that are specific for sequences in the organelle DNA (Shen & Fulthorpe, 2015). The DNA extraction is performed without distinction of host or bacterial DNA. The DNA extract is then subjected to enzymatic digestion (authors suggest 16 h at 37 °C) with enzymes that have restriction sites downstream and upstream of universal primer binding sites in chloroplast DNA. As before, a long enzymatic incubation step may allow for bacterial proliferation, affecting the original community structure. The digested samples are then used for PCR amplification using universal primers and the adequate sized bands are excised from the agarose gel after electrophoresis. This requirement to extract the band from the gel may result in loss of information during sample processing. The method suggested by these authors only aims at reducing host DNA contamination from chloroplast DNA. As such, mitochondrial DNA contamination is still observed after sequencing. Finally, despite the methodology applied, the authors still removed reads belonging to chloroplasts from the sequencing data.

(iii) During PCR amplification, using primers that perform a preferential amplification of bacterial DNA template (e.g., Chelius & Triplett, 2001), or using PNA blockers that prevent amplification of host DNA template (Lundberg et al., 2013).

Regarding the first approach, a number of primers that are less likely to lead to organelle DNA amplification have been described. By far, the most used is the 799f primer, which was firstly described by Chelius & Triplett (2001). This approach allows for a regular, non-differential DNA extraction, yielding host DNA and bacterial DNA in the total DNA pool. From this DNA pool, the amplification of the 16S rRNA gene from chloroplast DNA is excluded due to sequence mismatches with the primer, and the amplification of the 18S rRNA gene from mitochondrial DNA results in an amplicon with a distinct molecular weight from that obtained from bacterial 16S rRNA gene amplification. As a result, the bacterial amplicon can be excised from an agarose gel, and used for downstream applications such as sequencing (Chelius & Triplett, 2001). Nevertheless, this method presents some disadvantages. In addition to presenting sequence mismatches with the 16S rRNA gene sequence from chloroplast DNA, the primer 799f also presents mismatches with 16S rRNA gene sequences from bacterial groups, thereby introducing a structure and diversity bias into the resulting PCR product (Chelius & Triplett, 2001; Idris et al., 2004). Additionally, this method requires the extraction of a band from an agarose gel, which, as mentioned before, may lead to loss of information during sample processing. Furthermore, the method is not fully efficient, since authors have described having observed host DNA matches from sequencing data obtained with 454-pyrosequencing (e.g. Bodenhausen et al., 2013). Finally, the effectiveness of the 799f primer may result in minimal amplification of the bacterial sequences in the sample (Shen & Fulthorpe, 2015; present work, data not shown).

The use of PNA blockers also allows for a DNA extraction protocol that yields host and bacterial DNA in the final DNA pool. The PNA clamps block the amplification of host organelle DNA thus allowing for a PCR product enriched in bacterial sequences, without introducing the bias of bacterial group exclusion or of subsequent technical steps that are necessary with the use of differential primers (Lundberg et al., 2013). The PNA blockers have been successfully used before in Illumina-based studies of endophytes (Carrell et al., 2016). Shen & Fulthorpe (2015) also used PNA blockers and compared results to those obtained from enzymatic digestion of chloroplast template DNA. After 454-pyrosequencing data analysis, the authors either obtained similar results regarding read number and percentage of host DNA reads, or obtained more reads with the PNA blockers.

In the present work, the taxonomic attribution for the recovered OTUs from -PNA and +PNA samples revealed several instances of OTU affiliation to the 16S rRNA sequence from chloroplast DNA and to the 18S rRNA sequence from mitochondrial DNA. For the +PNA samples, removing these OTUs from the OTU table returned 393184 bacterial reads, which represents 22.32 % of the initial read number (Table 5.2). This suggests that the utilization of PNA blockers was not sufficient to block amplification of all host DNA competing templates, which could be due to (i) an insufficient quantity of PNA blockers in the reaction, (ii) non optimal PCR conditions for the PNA clamping, or (iii) due to mismatches in the template sequence.

(i) Regarding the quantity of PNA blockers used in the reaction, the concentration suggested in the manufacturer's instructions is 0.25  $\mu\text{M}$  per reaction, which could be increased up to 1  $\mu\text{M}$  to improve the clamping effect. After consulting with PNAbio representatives, it was decided to optimize the concentration of each blocker per reaction to 2  $\mu\text{M}$ . Nevertheless, it may not have been sufficient to block all the templates from healthy specimens of the species *H. portulacoides*. The number of copies of target template from host DNA is unknown in this halophyte, however, it is known that the number of copies of organelle genomes per cell can reach the tens of thousands in some plant species (Bendich, 1987).

(ii) Regarding the PCR conditions used for PNA clamping, the temperature and time of PNA annealing were performed as suggested by the manufacturer. These conditions should present enough stringency to ensure proper annealing of the PNA blockers with their respective targets. Optimization was performed on the mPNA and pPNA solubilization step prior to PCR amplification: the manufacturer's instructions suggest a 5 min incubation at 55 °C to solubilize the PNA blockers before adding to the PCR reaction. We performed this step according to their instructions and observed that the sequencing output still presented an unacceptably high amount of host DNA contamination (data not shown). After contacting the manufacturer, and per their suggestion, this step was optimized by increasing the incubation time to 30 minutes.

(iii) Regarding the possibility of mismatches between the PNA blockers and the target region of chloroplast and mitochondrial templates, it was demonstrated by Lundberg et al. (2013) that those targets are conserved among higher plants. We have checked *in silico* the sequence match of pPNA blockers against chloroplast sequences of a distant plant species (*Pinus pinaster*) and a closely related species (*Spinacea oleracea*) and observed that the match was 100 % in both cases (data not shown). As such, we assume that a plant species belonging to the same sub-family should not present relevant mismatches in such a conserved region of the chloroplast DNA. We

could not perform the same test for mitochondrial DNA since the database is lacking for this organelle's genome.

The use of mPNA and pPNA blockers in the PCR reaction is described as allowing for a higher yield of bacterial DNA amplification when host DNA is present as a competing template for the reaction (Lundberg et al., 2013). This was tested in the present work by sequencing a randomly chosen DNA extract (sample 11) amplified in the absence and presence of PNA blockers. As a result, a remarkable 976-fold increase of bacterial reads was observed (Table 5.3). The addition of PNA blockers did not prevent the amplification of the entirety of host DNA templates, as discussed above, however it allowed an immensely greater number of bacterial sequences to be analyzed in the sample.

**Table 5.3** Comparison of number of reads resulting from PCR amplification in the absence and presence of PNA blockers.

	Sample 11 -PNA	Sample 11 +PNA
Total number of reads	16229	32533
Database match		
Bacterial 16S rRNA	4 (0.025 %)	3905 (12.0 %)
Mitochondrial 18S rRNA	3394 (20.91 %)	7980 (29.53 %)
Chloroplast 16S rRNA	12826 (79.03 %)	20266 (62.29 %)

Other studies of endophytic bacterial diversity that apply Illumina-based sequencing technology usually do not mention which method, if any, was used to avoid host DNA contamination. Additionally, when OTUs that show affiliation to host DNA are removed from the sequencing data, the percentage of removed reads is often concealed. This lack of information in endophytic diversity reports represents a considerable obstacle in accurately comparing our results to those published to date. To the best of our knowledge, retaining roughly 22 % of the initial reads in this type of study is overall in agreement and/or superior to what is observed in the literature.

### **Sample analysis: alpha-diversity and rarefaction**

All 30 samples were analyzed for their alpha-diversity metrics, namely Shannon's diversity, Simpson's dominance (complementary and inverse dominance), Pielou's evenness, Chao's coverage estimators and richness determination (Table 5.4).

A wide range of values was observed throughout the 30 samples for Shannon's diversity, richness, Pielou's evenness and Simpson's indices. For most plant specimens (12 of the 15) Shannon's diversity ( $H$ ) was higher for BG tissues than for the respective AG tissues. The same pattern was observed for 13 of the 15 plant specimens in terms of richness. The highest  $H$  value (5.07) was observed for a BG sample (sample 16), and the lowest  $H$  value (1.32) was observed for an AG sample (sample 1). These highest and lowest  $H$  values were coincident for the samples with extreme values for Pielou's evenness (samples 16 and 1, respectively).

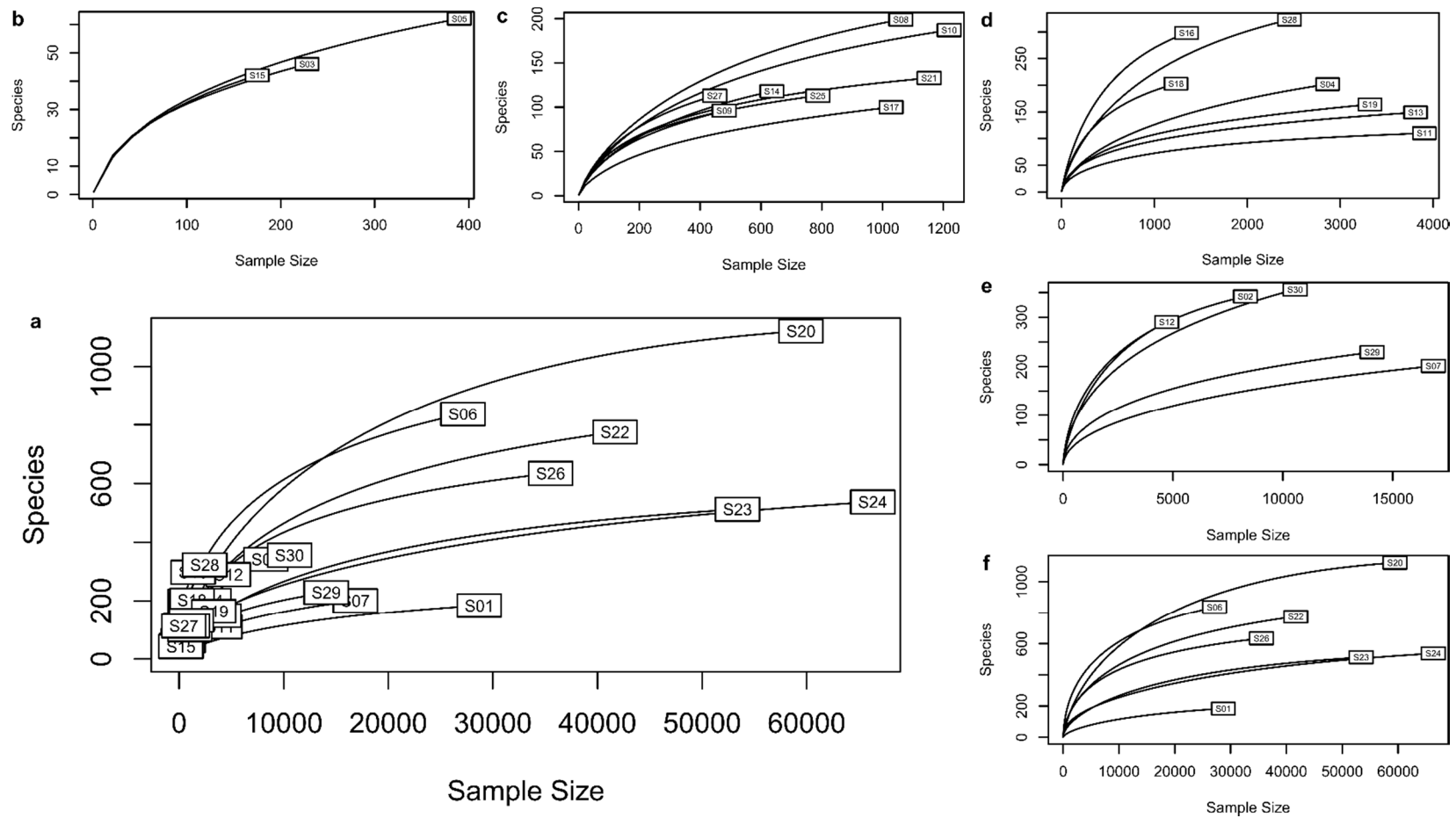
Rarefaction curves allow an assessment of the sequencing depth for all samples (Figure 5.2a). A plateau was observed for most samples, indicating that a good sequencing depth was obtained (Figure 5.2b-f). Chao's coverage estimation values for all samples ranged from 89.76 % to 99.79 % (Table 5.4). These results from plateaued rarefaction curves and elevated coverage values indicate that the communities for the different samples were well characterized, even in the presence of a prominent host DNA contamination in the sequencing output. Available literature on Illumina-based studies of endophytic bacterial communities shows either lower or comparable values for coverage estimation values (60 to 99.3 %; Table 5.5). Other alpha-diversity metrics were not compared with other studies since they are inherently highly dependent on the methodology, plant species, and other factors. Coverage estimation values were chosen for comparison to have an overview of what is achievable with Illumina-based studies of endophytes using samples from different tissues and plant species.

**Table 5.4** Alpha-diversity metrics results.

Sample	Shannon (H)	Richness (OTUs)	Alpha-diversity metrics			Chao's coverage
			Pielou (J)	Simpson (1-D)	Simpson (1/D)	
1	1.32	184	0.252	0.631	2.71	99.78
2	3.28	342	0.562	0.859	7.07	99.00
3	3.22	46	0.842	0.941	17.09	91.25
4	3.36	202	0.632	0.909	11.05	97.31
5	3.21	62	0.777	0.915	11.79	93.86
6	4.59	837	0.682	0.967	29.93	99.25
7	1.77	201	0.334	0.657	2.92	99.54
8	4.4	199	0.832	0.975	40.11	93.69
9	3.65	96	0.799	0.938	16.03	91.24
10	4.11	187	0.785	0.954	21.96	94.84
11	2.82	110	0.6	0.87	7.67	99.36
12	3.17	290	0.56	0.896	9.59	97.75
13	3.42	149	0.683	0.926	13.44	98.87
14	3.75	118	0.785	0.942	17.21	92.47
15	3.13	42	0.838	0.933	14.82	89.76
16	5.07	298	0.889	0.988	81.76	93.47
17	3.05	100	0.662	0.892	9.23	96.01
18	4.49	203	0.845	0.978	44.66	95.14
19	3.1	164	0.607	0.872	7.83	98.37
20	2.31	1122	0.33	0.622	2.64	99.76
21	4.06	133	0.83	0.97	33.38	96.79
22	3.66	775	0.551	0.928	13.86	99.52
23	2.51	512	0.402	0.798	4.94	99.78
24	3.35	537	0.533	0.919	12.37	99.79
25	3.96	113	0.837	0.962	26.44	95.43
26	3.59	634	0.557	0.921	12.68	99.61
27	4.12	113	0.871	0.972	36.07	90.2
28	4.45	323	0.771	0.966	29.78	95.8
29	2.24	229	0.412	0.67	3.03	99.43
30	3.57	356	0.608	0.937	15.84	98.83
Minimum	1.32	42	0.25	0.62	2.64	89.76
Median	3.39	200	0.67	0.93	13.65	97.53
Maximum	5.07	1122	0.89	0.99	81.76	99.79

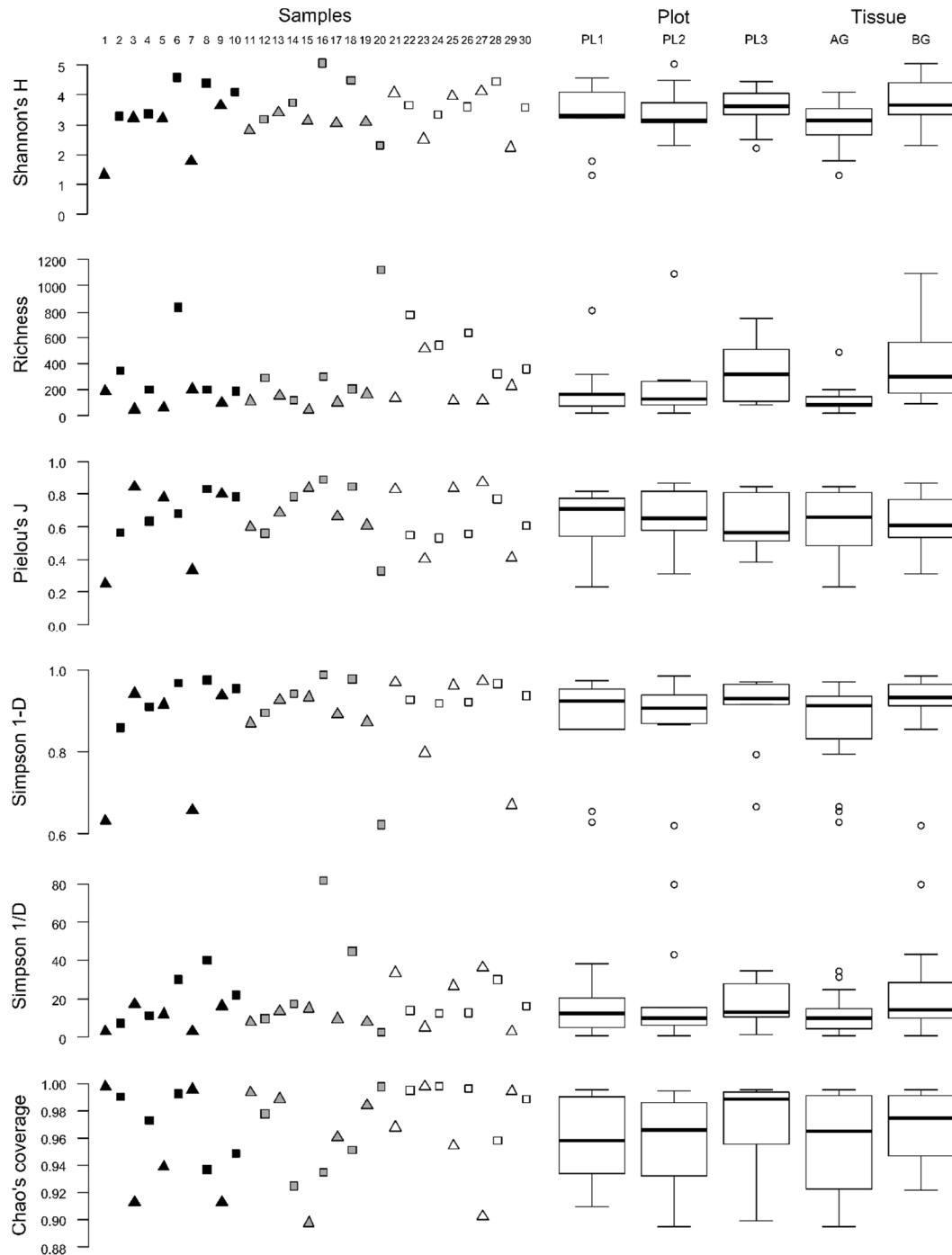
**Table 5.5** Coverage estimation values reported for Illumina-based studies of endophytic bacterial communities.

Plant species	Plant tissue and Coverage value (%)	Reference
<i>Oryza sativa</i>	Sprout: 84	Wang et al. (2016a)
	Stem: 60	
	Root: 83	
<i>Espeletia</i> sp.	Leaves: 96.6 to 99.3	Ruiz-Pérez et al. (2016)
Soya bean and Alfalfa	Soya bean nodules: 99.1	Xiao et al. (2017)
	Soya bean root: 96.6	
	Alfalfa nodules: 99.2	
	Alfalfa root: 97.3	



**Figure 5.2** Rarefaction curves for all samples. **a**, includes all samples; **b** to **f**, samples are separated according to sample size for easier visualization.

The alpha-diversity metrics were then analyzed according to sampling tissue and sampling plot, to assess if those were factors that significantly influenced the metrics (Figure 5.3).



**Figure 5.3** Alpha-diversity metrics results. Values are given for all 30 samples in the scatterplots (black: samples 1 to 10, plot 1; grey: samples 11 to 20, plot 2; white: samples 21 to 30, plot 3; triangles: aboveground tissues; squares: belowground tissues); and for the three plots (PL1, PL2 and PL3) and two tissues (AG and BG) in the boxplots.



Statistical analyses were performed on the alpha-diversity metrics. First, normality was assessed for each metric, and was only observed for Shannon's H (Shapiro-Wilk normality test:  $W = 0.97476$ ,  $p = 0.6756$ ). Consequently, the tests performed on the values for Shannon's H followed normality assumptions (ANOVA and Student's t-Test), while tests performed on all other metrics did not respect this assumption (Kruskal Wallis and Wilcoxon tests). We then tested if collecting the plants from different plots had an effect on the alpha-diversity metrics, and observed that there were no significant differences according to sampling plot.

Finally, we tested if the metrics varied according to the sampling tissue, and observed that there were significant differences for Shannon's H (BG > AG; Student's t-Test:  $t = -2.7905$ ,  $df = 27.481$ ,  $p = 0.009454$ ). Tissue-based differences were also observed in the culture-dependent approach for plants collected from the same site (Chapter 3; summary in Table 5.6). This pattern of differences in Shannon's diversity according to sampling tissue, where roots/BG tissues are more diverse than leaves/AG tissues, have also been reported in other Illumina-based studies of endophytic bacterial communities (Wang et al., 2016a; Xiao et al., 2017).

**Table 5.6** Comparison of average values of alpha-diversity metrics from culture-dependent and culture-independent methodologies.

Alpha-diversity metric	Culture-dependent (site C)	Culture-independent
Shannon's diversity (H)	AG: 2.08	AG: 3.04
	BG: 2.18	BG: 3.81
Pielou's evenness (J)	AG: 0.84	AG: 0.65
	BG: 0.90	BG: 0.66

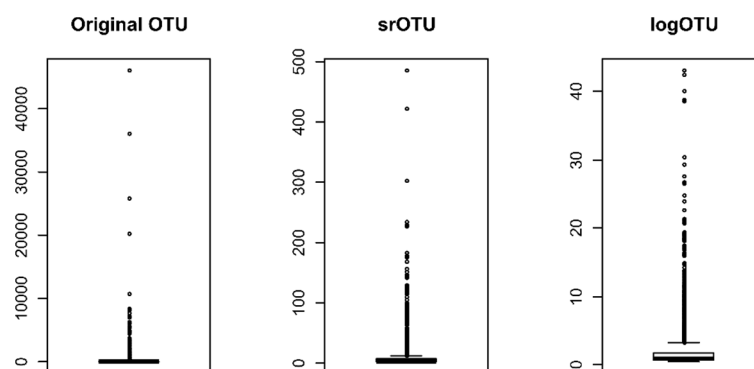
AG, aboveground tissues; BG, belowground tissues.

The analysis of richness also revealed statistical significance according to sampling tissue (BG > AG; Wilcoxon Rank Sum test:  $W = 22$ ,  $p = 0.0001889$ ). This has also been observed in Illumina-based studies of endophytic bacterial communities of *Aloe vera* (Akinsanya et al., 2015) and soya bean, but not for alfalfa (Xiao et al., 2017). Higher richness in BG tissues was not the pattern observed for the number of OTUs when the culture-dependent approach was used (Table 3.2). The difference in the methodologies used could justify this difference in results, since one of the major disadvantages of culture-dependent methods is that only a small fraction of the community is assessed (Amann et al., 1995), which, in this case, is likely a result of culture media not providing the necessary nutrients for all the bacteria present in a community (Tanaka et al., 2014).

### Community analysis: structure

The structure of the community was analyzed for all 30 samples. As observed above, the samples differ in read number (Table 5.2) and sequencing depth (Figure 5.2). When faced with such disparity in total read number per sample, some authors rely on sub-sampling of the original OTU table in order to obtain a representative OTU table with a minimal common read number for all samples. This method, however, results in loss of information and has been highly criticized (McMurdie & Holmes, 2014). In the case of the present work, applying this methodology would reduce sample sizes to roughly 200 reads per sample, which would exclude 98.47 % of the obtained bacterial reads. Since our data is intrinsically prone to loss of information due to host DNA contamination, this methodology will not be applied in the present work. As an alternative, the effect that the disparity in sample size presents can be reduced by using an adequate distance measure to build a resemblance matrix to further study the relationships among the samples (McMurdie & Holmes, 2013).

The original OTU table was subjected to data transformations in order to understand the effect of OTU abundance: (i) a square root transformation (srOTU) to down-weight the high-abundance OTUs; (ii) a logarithmic transformation (logOTU, with  $\log_{10}$ ) to enhance the importance given to low-abundance OTUs, and to lower the weight of high-abundance OTUs. The boxplots depicted in Figure 5.4 show the distribution of total number of reads per OTU for the original OTU table, and for the transformed logOTU and srOTU tables.



**Figure 5.4** Boxplot of total number of reads per OTU for the original and transformed OTU tables. srOTU, square root transformed OTU table; logOTU, logarithm ( $\log_{10}$ ) transformed OTU table.

The original OTUs and transformed OTUs tables were then used to create resemblance matrices using the Bray-Curtis distance measure. Statistical analyses of these matrices revealed that the tissue of origin (AG or BG) significantly affected the structure of the community (Adonis:  $p = 0.001$  for all OTU tables). However, the percentage of variation that this factor represented was relatively low (Adonis:  $R^2 = 8.69\%$ ,  $13.02\%$  and  $11.97\%$  for the original, logOTU and srOTU tables, respectively). These percentage values indicate that other variables are contributing for the majority of the variability observed among the samples (87 to 91 %).

Other factors that could contribute to the observed variability were then analyzed: plant, monospecific stand and plot of origin. The community structure was only significantly affected by plot of origin, and only when logarithmic or square root transformation was applied to the OTU tables (Adonis:  $p = 0.011$  and  $0.022$  for logOTU and srOTU, respectively). Nevertheless, the contribution of this factor was also low (Adonis:  $R^2 = 9.72\%$  and  $9.36\%$ , respectively) and a high percentage of variability was still left unaccounted for ( $> 77\%$ ). In both cases, the statistically different plots were Plot 2 and Plot 3 (Pairwise Adonis:  $p = 0.012$  and  $0.036$  for logOTU and srOTU, respectively). No interaction of factors revealed statistical significance. These results indicate that over 77 % of the variability observed is not accounted for by the analyzed factors. This residual variability may be due to:

- (i) Physical or chemical factors that were not analyzed in the present work. As discussed before (Chapter 3), factors such as soil water content, organic matter content, fine particle content, redox potential, pH, and conductivity should be assessed in order to better understand the effects of the environment on the community. However, since the fifteen plant samples at study were collected from one site, it is not expected that these factors should vary enough to impose significant changes. Furthermore, changes in such factors should be reflected by the 'monospecific stand' factor that was analyzed and found to not be significant in explaining the observed variability.
- (ii) Methodology-induced variability throughout sample handling steps. To assess if DNA extraction of the different samples, PCR amplification, or amplicon sequencing were the source of the observed variability, replication should be performed at all the different steps of the methodology.

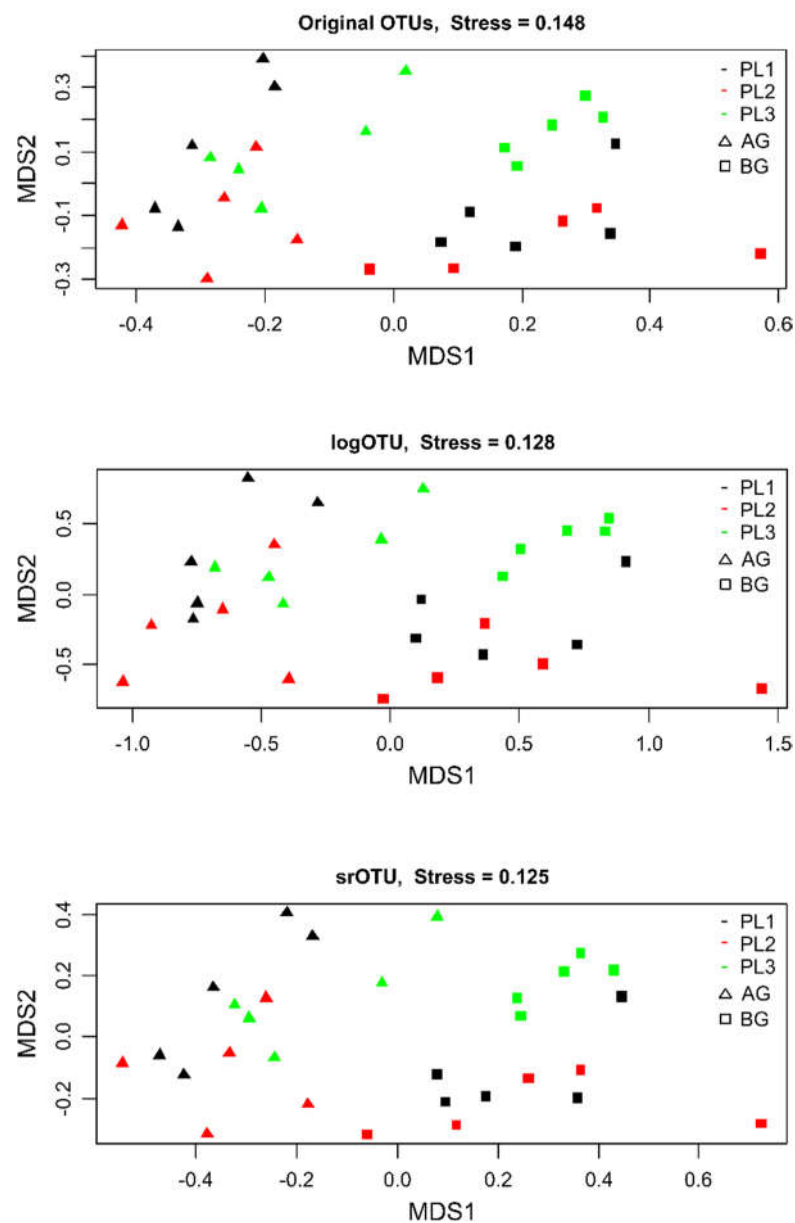
(iii) Bias in OTU formation introduced by read length. The raw reads were quality-filtered according to several parameters, such as read length, and reads with less than 150 bp were discarded. It is possible that reads that should be grouped into the same taxonomic unit are being separated due to a bias introduced by difference in read length. This would introduce additional diversity that does not correspond to the reality of the endophytic bacterial community. This possible bias could influence the low-abundance OTUs in the sense that not many of the OTUs would present a length that is compatible with OTU formation using the parameters that were used. In the case that this bias exists, it is wiser to look at the most abundant OTUs, even if it counters the aim of doing such an in-depth methodology like Illumina-based sequencing.

(iv) Heterogeneity in the community may be a reflection of transitory associations due to adaptation of the ever-changing surrounding biotic and abiotic factors (reviewed in Vandenkoornhuyse et al., 2015).

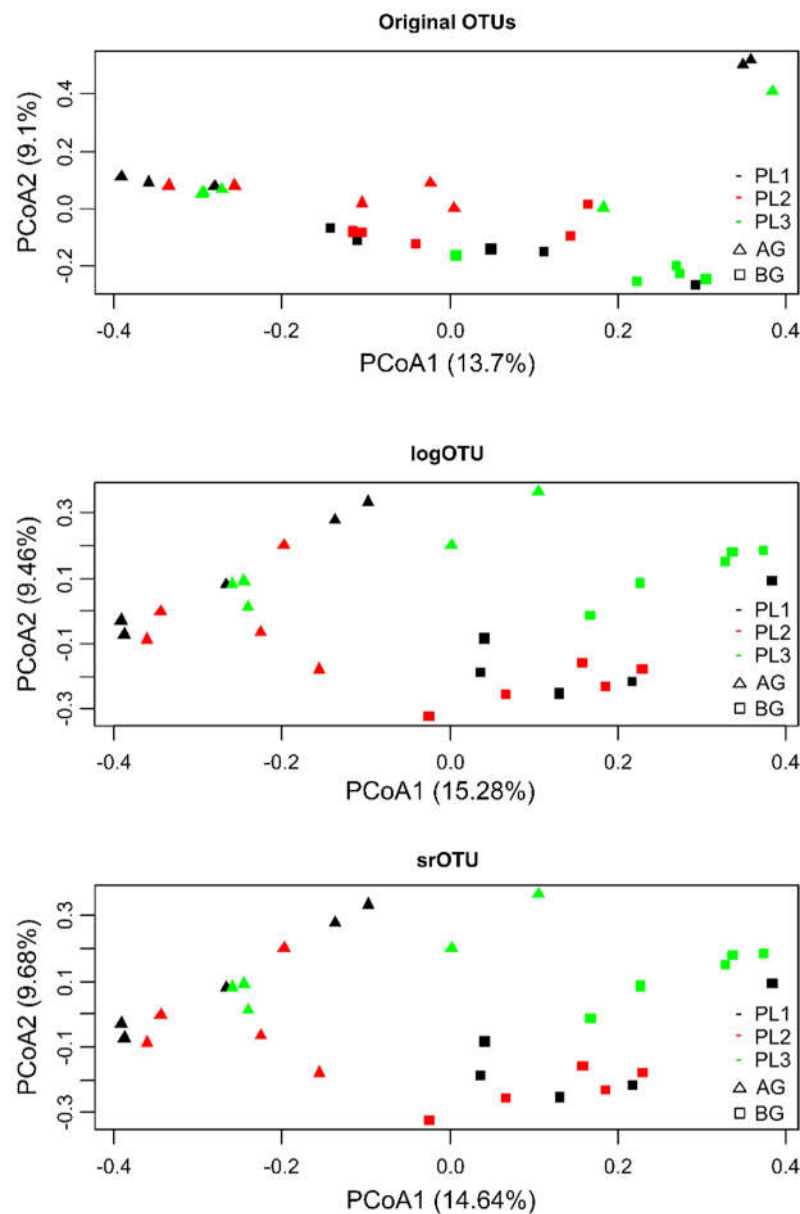
Ordination methods were then performed to assess the existence of patterns according to the factors that revealed a significant effect on the community structure: plot and tissue of origin. The ordination methods chosen were the rank-based non-metric multidimensional scaling (NMDS; Figure 5.5) and principal coordinates analysis (PCoA; Figure 5.6), and they were performed on the original and transformed OTU tables.

The NMDS representations revealed a sizeable structural variation among the 30 samples, which does not appear to be substantially influenced by the OTU table used as input. The NMDS representations revealed a clear pattern that separated the samples according to sampling tissue across the first axis (MDS1), which is in accordance with the statistical analysis performed above. A clear pattern is not observed for plot of origin: Plot 2 and Plot 3 appear to be more distinguishable from each other, but both have a coincident distribution with samples from Plot 1.

The stress values associated to each representation reflect the distortion applied to the relationships among samples, which is necessary to display the distances in a bi-dimensional environment. The stress values obtained are acceptable ( $< 0.2$ ), and are the lowest when the data is transformed by applying square root to the OTU abundance values, and the highest for non-transformed OTU values.



**Figure 5.5** Non-metric multidimensional scaling (NMDS) of the endophytic community according to plot and tissue of origin. PL1, sampling plot 1; PL2, sampling plot 2; PL3, sampling plot 3; AG, aboveground; BG, belowground.



**Figure 5.6** Principal coordinates analysis (PCoA) of the endophytic community according to plot and tissue of origin. PL1, sampling plot 1; PL2, sampling plot 2; PL3, sampling plot 3; AG, aboveground; BG, belowground.

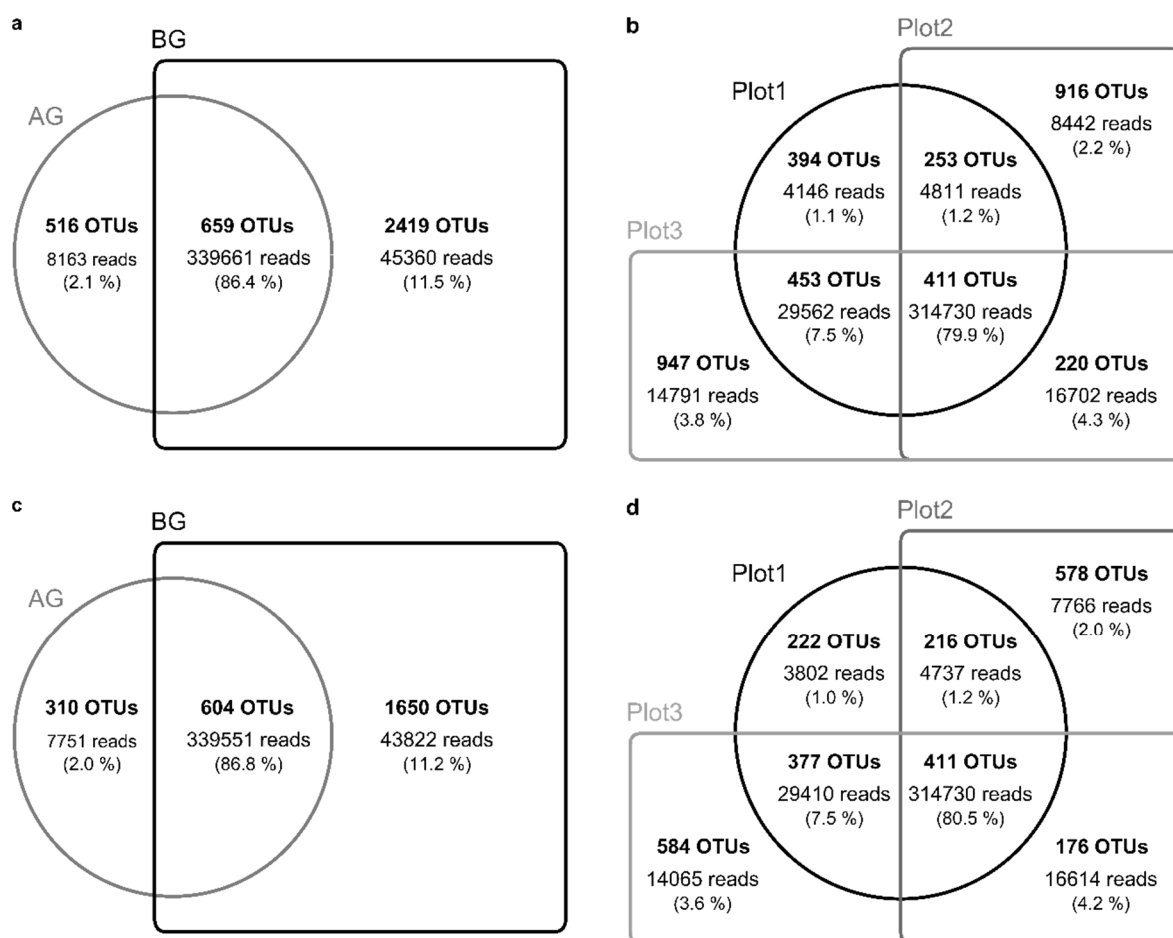
The PCoA representation is in accordance with the NMDS representation for the transformed data (logOTU and srOTU). It is noteworthy that, for the transformed data, the separation of the samples across the most explicative axis (PCoA1) reflects the factor tissue of origin, and accounts for 14.64 to 15.28 % of the variability observed among the samples in the bi-dimensional reduced environment. Again, this separation is in accordance with what was observed in the statistical analyses.

The PCoA representation for the original OTUs, however, tells a different story. The dispersion of the samples in the graphic representation is not as marked, especially across the second axis (PCoA2). This is not observed in the NMDS method because, since it is a non-metric approach, it uses rank-based distances, whilst the PCoA uses absolute distances. These results suggest that the high-abundance OTUs (which are down-weighted in the transformed OTU tables) are highly influential in determining the distances among the 30 samples. Additionally, no marked differences are observed for the different data transformations (logOTU vs srOTU), suggesting that the rare OTUs up-weighted in the  $\log_{10}$  transformation do not exert relevant influence on the relationships among the samples.

The most abundant OTUs influence the distribution of the samples in reduced space representations. This was observed when comparing (i) PCoA and NMDS representations for the non-transformed OTUs, and (ii) PCoA representations for the original and transformed OTU tables. There is statistical congruence showing that samples originated from AG tissues differ from those from BG tissues. Although it was not verified in all occurrences, the plot of origin should also be taken into account as a factor of OTU variability in the community.

Venn diagrams were then obtained to understand the proportion and abundance of OTUs that were exclusive and shared among tissues and plots (Figure 5.7). This is a different approach from the reduced space representations (NMDS and PCoA) since the Venn diagram does not take into account the distances among samples based on OTU abundance. The Venn diagrams here presented are useful to understand the OTU composition of the different levels of the factors tissue and plot. Venn diagrams were built including all OTUs (Figure 5.7a-b) and excluding very rare OTUs (less than 3 reads; Figure 5.7c-d) to understand their effect among the sampling tissues and plots.

The Venn diagrams built using all the OTUs found in the community (3593) revealed that the majority of the OTUs were found exclusively in BG tissues. This heterogeneity of OTU occurrence is in accordance with what was observed in the statistical analysis and also in the reduced space representations, where the sampling tissue was a significant factor in accounting for the differences among samples.



**Figure 5.7** Venn diagrams built using OTUs from the endophytic bacterial community. OTUs are grouped according to occurrence in sampling tissues (**a**, **c**) and plots (**b**, **d**). Diagrams **a** and **b** were built using all OTUs from the community, while diagrams **c** and **d** were built using only the OTUs with more than 2 reads. The corresponding read numbers and their percentages are also presented. AG, aboveground; BG, belowground.

The Venn diagram for sampling tissue showed that the number of reads exclusive in BG tissues did not reflect the OTU richness. In fact, the majority of bacterial reads were observed in the OTUs that were shared between AG and BG tissues. This suggests that the majority of the endophytic bacteria occurrences are shared among the tissues of the halophyte. As explored in Chapter 3, it is reported in the literature that endophytic bacteria colonize the plant through openings in the root system, and that these bacteria then colonize the inner tissues of the plant, reaching AG tissues as well (reviewed in Lodewyckx et al., 2002). The 516 AG tissue-exclusive OTUs may result from alternative entry points into the halophyte, as was also discussed in Chapter 3.



In the Venn diagram for OTU occurrence grouped according to sampling plot (Figure 5.7b), it can be observed that the majority of the OTUs (2083) are unique to one of the sampling plots. These plot-exclusive OTUs are, however, present in low read numbers and account for 10.16 % of the total number of reads. As was the case for the Venn diagram according to tissue (Figure 5.7a), the majority of the reads are shared among all the plots (80.05 %).

The results suggest that low-abundance OTUs are enhancing the differences among sampling plots or tissues when the data are interpreted in terms of OTUs. This is confirmed by the Venn diagrams built excluding very rare OTUs (less than 3 reads; Figure 5.7c-d), where the numbers of OTUs and reads in exclusive groups (either sampling plots or tissues) decrease. This decrease is more marked in the case of plot-exclusive OTUs, where over 60 % of plot-exclusive OTUs for plots 2 and 3 are removed. These very rare OTUs may be responsible for the observed significant differences obtained in the statistical analysis that revealed the sampling plot as a significant factor for the OTU composition of the community.

Our earlier statement that the variability observed may be due to the methodology used is also supported: since Illumina sequencing allows for an extensive and in-depth analysis of each sample, many rare OTUs will be found, and those are likely to differ among different samples, or even among different sequencing runs.

### Community analysis: overall composition

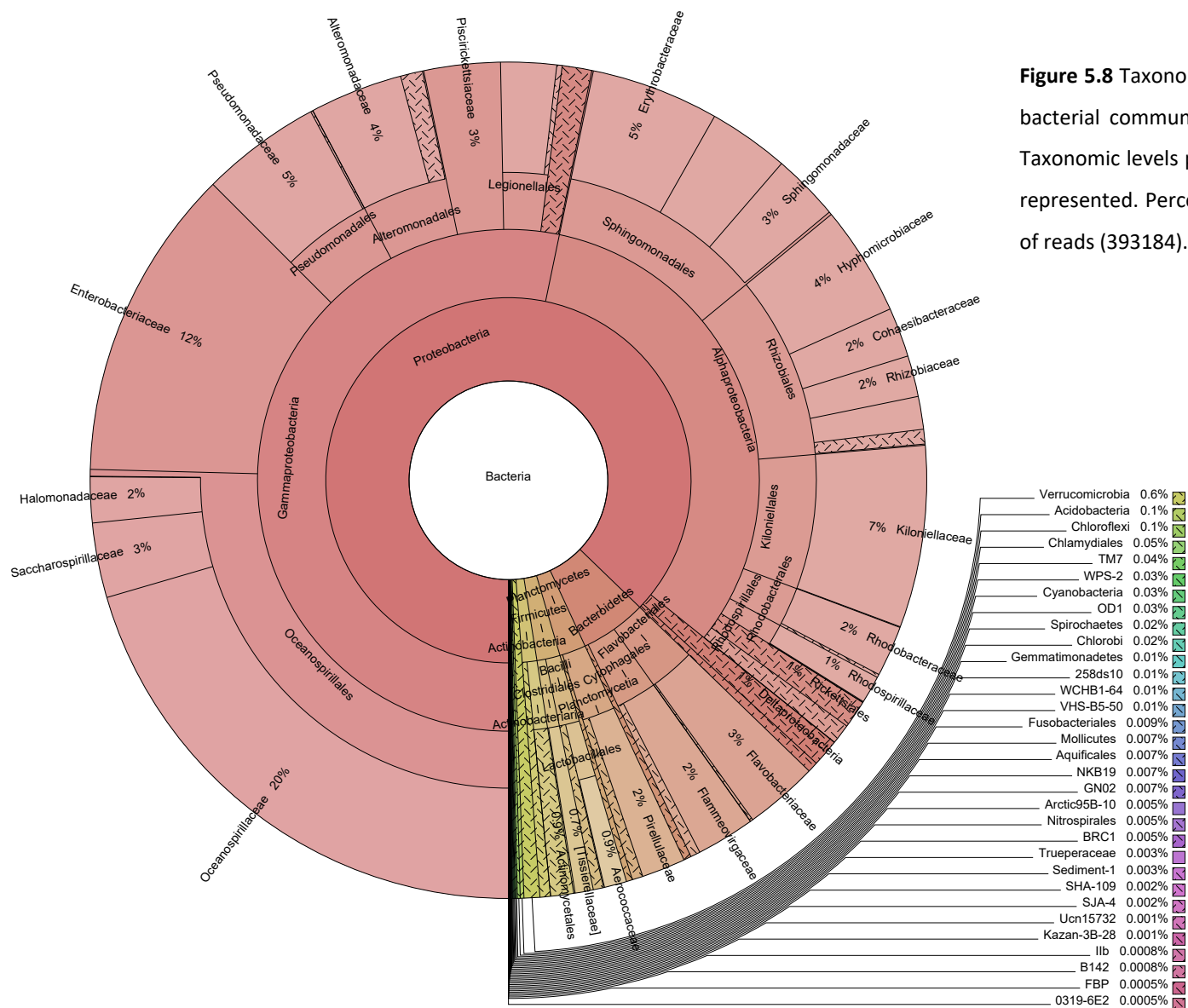
The taxonomic attribution of the endophytic bacterial community of *H. portulacoides* was assessed to understand the most relevant bacterial groups in the whole community and in the different sampling tissues. Sampling plot was also considered, but was less focused on.

Figure 5.8 depicts the taxa in the total community at the levels phylum, class, order and family. Overall, 37 phyla, 85 classes, 115 orders, 182 families and 232 genera were observed (detailed lists are presented in Annex II).

The dominant phylum observed in the endophytic bacterial community was Proteobacteria (62.5 % of OTUs; 87 % of reads). Other main phyla included Bacteroidetes (10.7 % of OTUs; 6 % of reads) and Planctomycetes (8.8% of OTUs; 2 % of reads). Firmicutes (2.4 % of OTUs; 2 % of reads) and Actinobacteria (5.2 % of OTUs; 1 % of reads) were also present. These results are in accordance with what was observed in the culturable fraction of the endophytic bacteria of this halophyte (Chapter 3), with the exception that the phylum Planctomycetes was not observed in the culture-based method.

The phyla of endophytic and other plant-associated bacteria observed in different halophyte species are listed in Table 5.7. All phyla observed in the other studies stated in Table 5.7 were also detected in the present work. The phyla Proteobacteria, Actinobacteria, Firmicutes and Bacteroidetes appear to be, overall, very frequent in the plant-associated bacterial communities among the different halophyte plant species, even in those not closely related to *H. portulacoides*. These four phyla have also been described in association with other plants (Bulgarelli et al., 2013). Previous reports of bacteria associated with *H. portulacoides* also reported the presence of Proteobacteria, Actinobacteria and Bacteroidetes (Oliveira et al., 2014a; Oliveira et al., 2014b).

The phylum Planctomycetes occurs in a considerable percentage in our Illumina-based analysis (approximately 9 %), however its detection has not been very frequent among the communities of halophytes. It was detected in the rhizosphere of *Arthrocnemum macrostachyum* by 454-pyrosequencing (Mora-Ruiz et al., 2016) and in the sediment colonized by *Spartina alterniflora* in the analysis of a clone library (Shuang et al., 2009). Our report is, to the best of our knowledge, the first of Planctomycetes as an endophyte of halophytes. Our previous effort of analyzing the endophytic community of *H. portulacoides* using a culture-based approach did not reveal any isolate from the phylum Planctomycetes (Chapter 3, Table 5.7).



**Table 5.7** Phyla of plant-associated bacteria observed in different halophyte species.

Plant species	Taxonomic relation	Sample	Method	Phylum									Reference
				Proteobacteria	Actinobacteria	Firmicutes	Bacteroidetes	Planctomycetes	Acidobacteria	Cyanobacteria	Spirochaetes	Chloroflexi	
<i>H. portulacoides</i>	--	Endophytes	Illumina	63%	5%	2%	11%	9%	✓	✓	✓	✓	Present Chapter Chapter 3
<i>H. portulacoides</i>	--	Endophytes	Culturable	64%	23%	10%	3%	-	-	-	-	-	
<i>H. portulacoides</i>	--	Endophytes	Culturable	✓	✓	-	-	-	-	-	-	-	Oliveira et al. (2014a)
<i>H. portulacoides</i>	--	Sediment and rhizosphere	454-pyrosequencing	60%	✓	-	8%	-	-	-	-	-	Oliveira et al. (2014b)
<i>Spartina alterniflora</i>	Not closely related	Sediment	Clone library	37%	✓	✓	30%	✓	✓	-	-	-	Shuang et al. (2009)
<i>Salicornia</i> sp.	Same family	Sediment and rhizosphere	DGGE and culturable	✓	✓	✓	✓	-	-	-	-	-	Mapelli et al. (2013)
Salicornioideae halophytes	Same family	Leaf endophytes and epiphytes	454-pyrosequencing and culturable	92%	1%	6%	1%	-	✓	-	-	-	Mora-Ruiz et al. (2015)
<i>Spartina alterniflora</i>	Not closely related	Root endophytes, rhizosphere and sediment	Illumina	45%	-	4%	11%	-	-	24%	4%	3%*	Hong et al. (2016)
<i>Spartina alterniflora</i>	Not closely related	Root endophytes, rhizosphere and sediment	Illumina	44%	-	12%	11%	-	-	-	-	4%	Su et al. (2016)
<i>Arthrocnemum macrostachyum</i>	Same family	Endophytes and rhizosphere	454-pyrosequencing	87%	-	9%	2%	2%*	✓	-	-	-	Mora-Ruiz et al. (2016)
<i>Salicornia europaea</i>	Same family	Root endophytes	454-pyrosequencing	95%	1%	0.6%	3%	-	-	-	-	-	Zhao et al. (2016)
<i>Salicornia europaea</i>	Same family	Endophytes and rhizosphere	Culturable	✓	✓	✓	-	-	-	-	-	-	Szymańska et al. (2016a)
<i>Aster tripolium</i>	Not closely related	Endophytes and rhizosphere	Culturable	✓	✓	✓	-	-	-	-	-	-	Szymańska et al. (2016b)

✓, present in the community; -, not detected in the community; \*, only in the rhizosphere.

The absence of Planctomycetes in the culturable fraction could be due to the fact that (i) bacteria from this phylum usually present a slow growth rate and can be outgrown by other bacteria that have faster growth rates; (ii) inadequacy of the culture media since the successful isolation of these bacteria is enhanced by the presence of vitamin solutions, micro- and macronutrients; (iii) a surplus of yeast extract or peptone, usually observed in rich culture media, may inhibit their growth (Lage & Bondoso, 2011; Lage & Bondoso, 2012). The specific nutritional requirements of this group of bacteria may justify their presence in the endosphere of a halophyte. According to Musat et al. (2006), Planctomycetes are common in marine sediments that are subjected to tides, which is an environment comparable to that of a salt marsh. The ecological role of this phylum is hypothesized to be related to providing competition with undesired microorganisms, avoiding dissection, participating in the marine nitrogen cycle by anaerobic ammonium oxidation, and potentially producing relevant metabolites (Lage et al., 2011).

The main classes observed in the endophytic community included Alphaproteobacteria (28.8 % of OTUs; 33 % of reads), Gammaproteobacteria (28.3 % of OTUs, 53 % of reads), Planctomycetia (7.2 % of OTUs, 2 % of reads), Flavobacteriia (5.3 % of OTUs, 3 % of reads), Deltaproteobacteria (3.7 % of OTUs, 1 % of reads) and Cytophagia (3.1 % of OTUs; 3 % of reads). Both Alpha- and Gammaproteobacteria were also found dominant classes when using the culture-based methodology applied in this report (Chapter 3).

The main families in number of OTUs in the bacterial endophytic community (comprising approximately 55 % of the OTUs) included *Rhodobacteraceae* (206 OTUs), *Enterobacteriaceae* (205), *Oceanospirillaceae* (192), *Pirellulaceae* (157), *Flavobacteriaceae* (155), *Alteromonadaceae* (147), *Erythrobacteraceae* (124), *Hyphomicrobiaceae* (114), *Sphingomonadaceae* (100), *Planctomycetaceae* (94), *Flammeovirgaceae* (74), *Pseudomonadaceae* (66), *Rhizobiaceae* (56), *Verrucomicrobiaceae* (51), *Rhodospirillaceae* (51), *Halomonadaceae* (48), *Piscirickettsiaceae* (47), and *Kiloniellaceae* (45). The families with higher number of reads (over 10000) were: *Oceanospirillaceae* (80518 reads), *Enterobacteriaceae* (47558), *Kiloniellaceae* (27656), *Erythrobacteraceae* (19438), *Pseudomonadaceae* (18298), *Hyphomicrobiaceae* (16712), *Alteromonadaceae* (14134), *Piscirickettsiaceae* (11587), *Saccharospirillaceae* (11358), *Flavobacteriaceae* (11164) and *Sphingomonadaceae* (10205). Several of the families found in the present work were also detected in diversity assays in closely related plants (same family; Table

5.8). The families *Enterobacteriaceae* and *Halomonadaceae* were frequently found in those studies: 4 out of 5 analyzed studies.

**Table 5.8** Families of plant-associated bacteria observed in closely related halophyte species. The bold and underlined prints indicate families in the present work with higher number of OTUs and reads, respectively.

Family	Mapelli et al. (2013)	Mora-Ruiz et al. (2015)	Mora-Ruiz et al. (2016)	Zhao et al. (2016)	Szymańska et al. (2016a)
<b><u>Alteromonadaceae</u></b>	-	✓	-	-	-
<i>Bacillaceae</i>	✓	✓	-	-	✓
<i>Bradyrhizobiaceae</i>	-	-	✓	-	-
<i>Brevibacteriaceae</i>	-	✓	-	-	-
<i>Chitinophagaceae</i>	-	✓	-	-	-
<i>Comamonadaceae</i>	-	✓	-	-	-
<b><u>Enterobacteriaceae</u></b>	-	✓	✓	✓	✓
<b><u>Flavobacteriaceae</u></b>	-	✓	✓	-	-
<b><u>Halomonadaceae</u></b>	✓	✓	-	✓	✓
<i>Methylobacteriaceae</i>	-	✓	-	-	-
<i>Microbacteriaceae</i>	-	-	-	-	✓
<i>Micrococcaceae</i>	✓	-	-	-	-
<i>Moraxellaceae</i>	-	-	✓	✓	-
<i>Oxalobacteraceae</i>	-	✓	-	-	-
<i>Planococcaceae</i>	-	-	✓	-	-
<b><u>Pseudomonadaceae</u></b>	-	✓	-	✓	✓
<b><u>Rhizobiaceae</u></b>	-	✓	-	-	-
<b><u>Rhodobacteraceae</u></b>	-	✓	-	-	-
<i>Salinisphaeraceae</i>	-	✓	-	-	-
<i>Sinobacteraceae</i>	-	✓	-	-	-
<b><u>Sphingomonadaceae</u></b>	-	✓	✓	-	-
<i>Staphylococcaceae</i>	-	✓	-	-	-
<i>Xanthomonadaceae</i>	-	✓	-	-	-

Twelve of the main families observed here were also observed in the culturable collection: *Alteromonadaceae* (3 isolates), *Erythrobacteraceae* (78), *Flammeovirgaceae* (1), *Flavobacteriaceae* (15), *Halomonadaceae* (100), *Hyphomicrobiaceae* (22), *Oceanospirillaceae* (3), *Pseudomonadaceae* (15), *Rhizobiaceae* (12), *Rhodobacteraceae* (41), *Rhodospirillaceae* (6), *Saccharospirillaceae* (1) and *Sphingomonadaceae* (19). Curiously, the family *Dietziaceae*, which was present in the culturable fraction with 19 isolates, was not detected with the Illumina-based approach.

Isolates affiliated to the families *Erythrobacteraceae*, *Aerococcaceae*, *Microbacteriaceae*, *Halomonadaceae* and *Vibrionaceae* tested positive for several enzymatic assays and plant growth promotion traits (PGP, Chapter 3). In the Illumina-based approach, these families were found in more or less abundant amounts: 19438 reads (124 OTUs) for *Erythrobacteraceae*, 3399 reads (17

OTUs) for *Aerococcaceae*, 60 reads (6 OTUs) for *Microbacteriaceae*, 6947 reads (48 OTUs) for *Halomonadaceae* and 1118 reads (21 OTUs) for *Vibrionaceae*.

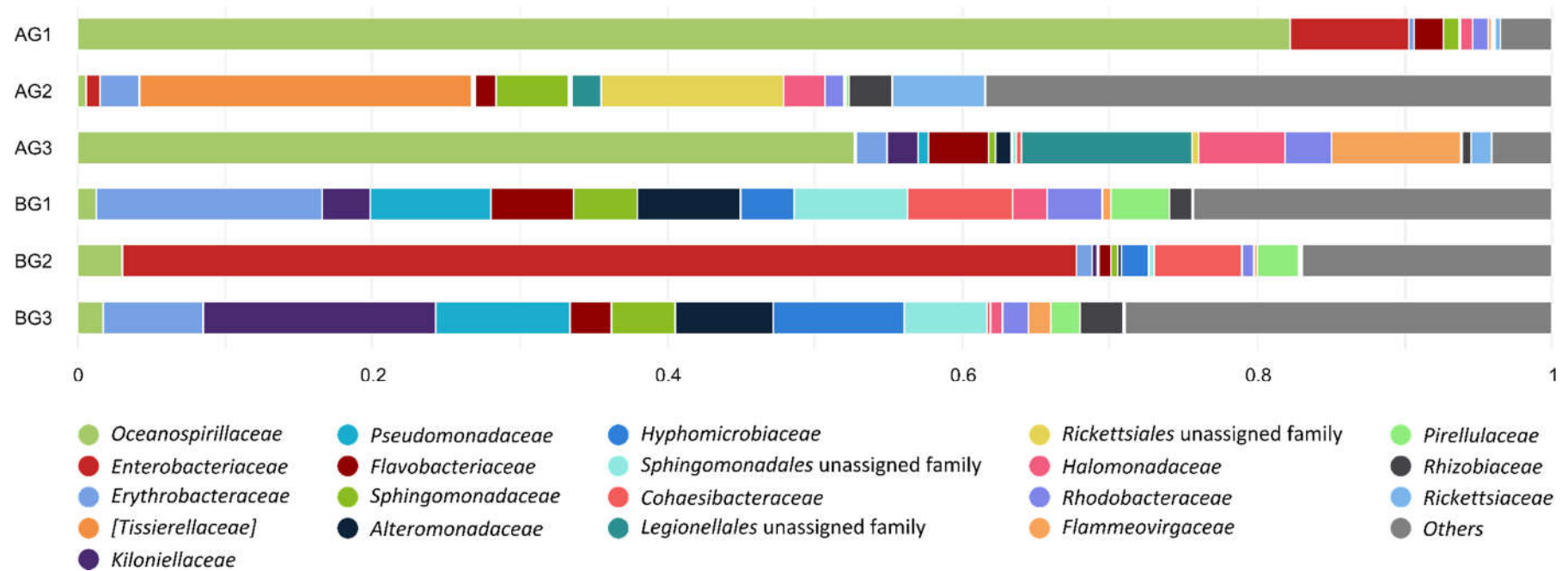
### Community analysis: core composition

The sampling tissue was determined to be a significant factor in the relationships between the samples of the endophytic community of *H. portulacoides*, and those differences were reflected in the NMDS and PCoA analyses. The sampling plot was also found to play a role in those relationships and, consequently, these two factors are now being analyzed for their influence in the taxonomic composition of the endophytic community from *H. portulacoides*.

Figure 5.9 shows the distribution of the twenty most read-abundant families across the different combinations of sampling plots and tissues. It is clear that there are differences in composition of the most dominant families between sampling tissues and sampling plots. A notable difference is observed for the combination of both tissues for Plot 2, where a clear dominance of [*Tissierellaceae*] and *Enterobacteriaceae* families are observed for AG and BG tissues, respectively.

Upon observation of the original data, it was clear that most of the reads from this family were observed in this plot in one single sample (99.5 % for [*Tissierellaceae*] in sample 11, and 91.4 % for *Enterobacteriaceae* in sample 20). This result contributes to explain why the sampling plot was determined to be a significant factor for the differences observed among the samples, since such a marked dominance is observed for these particular sets of OTUs.

In the culturable approach, the genera *Hoeflea*, *Labrenzia* and *Microbacterium* were exclusively found in BG tissues. With the Illumina-based approach, OTUs attributed to the family *Phyllobacteriaceae* (to which the genus *Hoeflea* is assigned) were found to either be exclusive to BG tissues (16 OTUs, 179 reads) or shared between both tissues (9 OTUs, 1472 reads). In the case of the family *Hyphomicrobiaceae* (to which the genus *Labrenzia* belongs), tissue-exclusivity was not observed, however less OTUs were found to be exclusive to AG tissues (3 OTUs, 19 reads), and more exclusive to BG tissues (100 OTUs, 2277 reads) and shared between both tissues (11 OTUs, 14416 reads). OTUs attributed to the family *Microbacteriaceae* (to which the genus *Microbacterium* belongs) were observed exclusive to AG tissues (2 OTUs, 16 reads), exclusive to BG tissues (1 OTU, 2 reads), and shared between both tissues (3 OTUs, 42 reads).



**Figure 5.9** Read abundance for the twenty top families according to sampling tissue and plot. AG, aboveground tissues; BG, belowground tissues; numbers refer to the sampling plot.



With the Illumina-based sequencing, several families were exclusively associated to one of the sampling tissues (15 to AG and 49 to BG tissues). For most of these occurrences, the read number was very low, which can justify their exclusiveness.

The tissue-exclusive OTUs with at least 10 reads were, for AG tissues: *Actinomycetaceae* (1 OTU, 14 reads), *Carnobacteriaceae* (1 OTU, 14 reads), *Leptotrichiaceae* (1 OTU, 32 reads), *Porphyromonadaceae* (3 OTUs, 139 reads), *Prevotellaceae* (3 OTUs, 62 reads), *Sanguibacteraceae* (1 OTU, 16 reads) and *Veillonellaceae* (4 OTUs, 117 reads).

The BG-exclusive families were: *[Balneolaceae]* (3 OTUs, 19 reads), *[Cerasicoccaceae]* (5 OTUs, 162 reads), *[Chthoniobacteraceae]* (10 OTUs, 103 reads), *Acetobacteraceae* (2 OTUs, 31 reads), *Alcanivoracaceae* (3 OTUs, 21 reads), *Arctic95B-10* (3 OTUs, 20 reads), *Desulfovibrionaceae* (4 OTUs, 33 reads), *Francisellaceae* (1 OTU, 27 reads), *Gemmataceae* (3 OTUs, 21 reads), *Holosporaceae* (1 OTU, 277 reads), *JdFBGBact* (1 OTU, 11 reads), *koll13* (3 OTUs, 56 reads), *Kordiimonadaceae* (2 OTUs, 20 reads), *Leptospiraceae* (5 OTUs, 28 reads), *Methylophilaceae* (3 OTUs, 197 reads), *Nitrospiraceae* (1 OTU, 10 reads), *Oleiphilaceae* (2 OTUs, 12 reads), *OM60* (11 OTUs, 396 reads), *Opitutaceae* (9 OTUs, 58 reads), *Parachlamydiaceae* (2 OTUs, 38 OTUs), *Phycisphaeraceae* (9 OTUs, 67 reads), *Polyangiaceae* (1 OTU, 32 reads), *Pseudonocardiaceae* (8 OTUs, 259 reads), *RB40* (6 OTUs, 42 reads), *Rhodobiaceae* (2 OTUs, 17 reads), *Sinobacteraceae* (5 OTUs, 52 reads), *SJA-101* (2 OTUs, 29 reads) and *Trueperaceae* (4 OTUs, 13 reads).

To achieve our goal of determining the core endophytome of *H. portulacoides*, we analyzed the AG and BG tissues separately in terms of taxonomic composition, and looked exclusively at the OTUs that are shared among all sampled plots, i.e., the core for each tissue of the halophyte. Our definition of core microbiome in the present work is the membership core, where the taxa shared among all the microbiomes of each plot are considered the core (Shade & Handelsman, 2011). Since fifteen samples of each tissue were sampled and are under analysis, the core endophytome further discussed should be representative of the population at study (Vandenkoornhuyse et al., 2015).

Removing the plot-specific OTUs still conserved approximately 75 % of the total reads obtained in the output. This indicates that the obtained core is substantial across the complex endophytic community. The core endophytome for each sampling tissue will be discussed below, alongside the possible ecological functions associated to those core taxa.

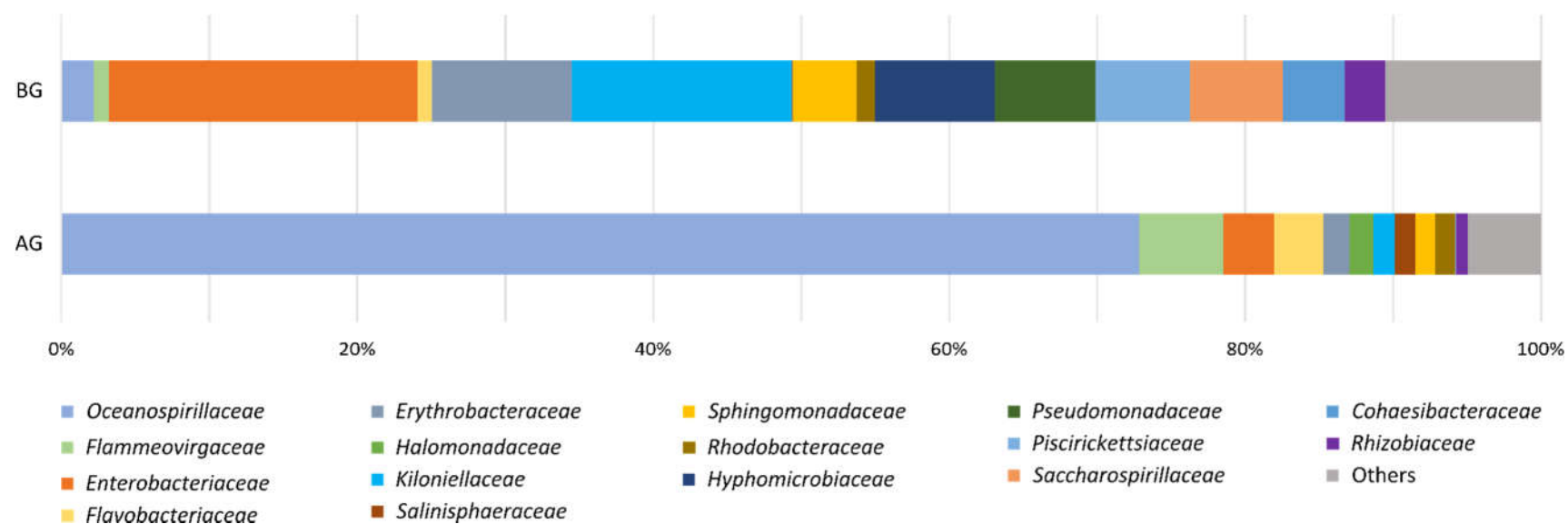
Although function can differ according to bacterial species or even strain (Glick, 2015), it is relevant to address the possible ecological functions, since the core endophytome is important for the plant holobiont (reviewed in Vandenkoornhuyse et al., 2015).

In the AG tissues, the core bacterial endophytome is composed of 104 OTUs, and the core BG bacterial endophytome is composed of 299 OTUs. Figure 5.10 shows the distribution of the most read-abundant families for the core of each sampling tissue.

The core endophytome of the AG tissues is overwhelmingly dominated by OTUs affiliated to the family *Oceanospirillaceae* (3 *Marinomonas* OTUs; 72086 reads) that comprises 72 % of the reads in the core endophytome for AG tissues. The low number of OTUs and large number of reads associated to the family *Oceanospirillaceae* suggest that well-adapted endophytes are dominant in the AG tissues of *H. portulacoides*. Regarding its possible ecological role in the endosphere, a metagenomics approach linked the family *Oceanospirillaceae* (genus *Marinomonas*) to functions such as the production of B12 vitamin, which is used in metabolism of plants and bacteria (such as Planctomycetes, as discussed above), pyruvate synthesis, response to environmental stress and detoxification, participation in nitrogen (e.g., ammonium assimilation and glutamine production), sulfur, phosphorus and carbohydrate metabolism, and iron uptake (Delmont et al., 2015).

The genus *Marinomonas* contains 24 species described thus far, presents worldwide distribution in marine environments, and has been isolated from the endosphere of the halophyte *Spartina maritima* (Lucena et al., 2016). Most species from this genus require salt for growth, and some can grow in the presence of 15 % salinity (Lucena et al., 2016). Traits observed in the endophytic community of *Spartina maritima* included enzymatic activities (protease, amilase, cellulase and lipase) and PGP traits (ACC deaminase activity, IAA and siderophore production, nitrogen fixation and phosphate solubilization; Mesa et al., 2015).

In the culturable fraction from the present work (Chapter 3), isolates belonging to the genus *Marinomonas* presented 2 to 7 enzymatic activities (cellulolytic and xylanolytic present in all), and PGP traits such as IAA production, ACC deaminase activity and presence of the *nifH* gene.



**Figure 5.10** Read abundance for the top core families for each tissue. AG, aboveground tissues; BG, belowground tissues.

Other families observed in the AG core endophytome included *Flammeovirgaceae* (2 unassigned OTUs; 5652 reads), *Enterobacteriaceae* (2 OTUs assigned to *Erwinia* and *Klebsiella*; 3387 reads), *Flavobacteriaceae* (6 OTUs attributed to *Leeuwenhoekiella*, *Mesonina* and *Salegentibacter*; 3283 reads), *Erythrobacteraceae* (11 OTUs attributed to *Citromicrobium*, *Lutibacterium* and unassigned genera; 1750 reads), *Halomonadaceae* (7 OTUs assigned to *Kushneria* and *Halomonas*; 1594 reads), *Kiloniellaceae* (1 *Thalassospira* OTU; 1428 reads), *Salinisphaeraceae* (4 *Salinisphaera* OTUs; 1379 reads), *Sphingomonadaceae* (7 OTUs attributed to *Novosphingobium*, *Sphingomonas* and unassigned genera; 1336 reads) and *Rhodobacteraceae* (8 OTUs attributed to *Paracoccus*, *Sulfitobacter* and unassigned genera; 1332 reads), which overall comprise 21 % of the OTUs from the AG core endophytome.

Isolates obtained from these families in the culturable approach (attributed to the genera *Leeuwenhoekiella*, *Mesonina*, *Citromicrobium*, *Kushneria*, *Halomonas*, *Thalassospira*, *Novosphingobium*, *Paracoccus* and *Sulfitobacter*) presented up to 6 (one *Halomonadaceae* isolate) enzymatic activities and PGP traits such as phosphate solubilization (only *Halomonadaceae*), IAA production (in all families), siderophore production (*Halomonadaceae*, *Erythrobacteraceae*, *Rhodobacteraceae* and *Kiloniellaceae*), presence of *nifH* gene (*Erythrobacteraceae*, *Halomonadaceae* and *Rhodobacteraceae*) and ACC deaminase activity (*Halomonadaceae*, *Erythrobacteraceae*, *Kiloniellaceae* and *Rhodobacteraceae*).

The core endophytome from the BG tissues was mainly populated by *Enterobacteriaceae* (3 OTUs, attributed to *Citrobacter*, *Erwinia*, and an unassigned genus; 35965 reads) and *Kiloniellaceae* (4 OTUs, 3 attributed to *Thalassospira* and one unassigned genus; 25657 reads), comprising 32 % of the reads in the BG core endophytome.

Genera belonging to the family *Enterobacteriaceae* have been widely found in plant-associated environments, including the endosphere (Hardoim et al., 2015). In the isolate collection this family was represented only by *Rahnella* spp. isolates, which presented proteolytic, cellulolytic and xylanolytic activities, and were able to solubilize phosphate, produce siderophores and IAA in high amounts (105.9  $\mu\text{g mL}^{-1}$ ), and presented ACC deaminase activity. The family *Enterobacteriaceae* contains many genera capable of promoting the growth of plants and whose strains present enzymatic activities and traits that are of interest for direct and indirect promotion of plant growth. Literature shows that the most observed traits are siderophore production, IAA production, and

phosphate solubilization. More details are now discussed for some genera of interest, namely *Citrobacter*, *Pantoea*, *Enterobacter* and *Erwinia*.

*Citrobacter*, one of the genera to which OTUs were assigned in the Illumina-based sequencing, has been associated with production of ammonium and IAA (Thomas & Upreti, 2016), ACC deaminase activity, phosphate solubilization, and siderophore production (Gontia-Mishra et al., 2017). In these studies, authors also observed that the *Citrobacter* spp. isolates were able to promote growth of different plant species: tomato (Thomas & Upreti, 2016) and wheat (Gontia-Mishra et al., 2017).

*Pantoea* was firstly detected as a plant pathogen, but has recently been found to also have a plant promotion effect, especially with a newly described species, *Pantoea alhagi* (Chen et al., 2017). Plant growth promotion has been observed in wheat, *Arabidopsis thaliana* (Chen et al., 2017) and tomato (Thomas & Upreti, 2016). Detected traits included the production of ammonium and IAA, phosphate solubilization (Thomas & Upreti, 2016), nitrate reduction, siderophore production, exopolysaccharides (EPS), and cellulolytic and proteolytic activities.

*Enterobacter*, alongside *Pantoea*, comprise members that present a range of interactions with plants from pathogenicity to mutualism (Hardoim et al., 2015). Member of the genus *Enterobacter* have been thoroughly explored for their abilities as a plant growth promoters, and have been found to promote growth of different plant species, such as tomato (Thomas & Upreti, 2016), wheat (Gontia-Mishra et al., 2017) and coconut seedlings (Gupta et al., 2014). Several traits have been associated with the genus *Enterobacter*, such as production of ammonium and IAA (in some cases with high levels; Ribeiro & Cardoso, 2012), phosphate solubilization, nitrogen fixation, siderophore production (Thomas & Upreti, 2016; Ribeiro & Cardoso, 2012), ACC deaminase activity, ammonium and EPS production, and cellulolytic, proteolytic and lipolytic activities (Gontia-Mishra et al., 2017). An analysis of the genomes of *Enterobacter* strains revealed additional genes of interest for the application of plant growth promotion, including genes that allow an enhancement of nutrient availability, decrease in pathogenic fungi, resistance to oxidative stress, quorum sensing, breakdown of aromatic and toxic compounds and other abiotic stressors. Other genes present in the genomes included those related to ACC deaminase activity, IAA and siderophore production, phosphate solubilization, H<sub>2</sub>S production, peroxidases, chitinases and catalases (Gupta et al., 2014). Isolates from maize belonging to the genera *Enterobacter* and *Pantoea* have been explored for their potential as biofertilizers in tropical soils due to their ability to solubilize phosphate (de Abreu et al., 2017).

Finally, the genus *Erwinia*, which was also assigned to our endophytic OTUs in the Illumina-based sequencing, has been mainly isolated from plant tissues where it has shown both pathogenic and non-pathogenic behavior, depending on the bacterial species. Genome-based evidence showed that these differences among species (namely the phytopathogenic *E. amylovora* and *E. pyrifoliae* and the epiphytic *E. tasmaniensis* and *E. billingiae*) are due to recombination events. Additionally, the same study suggested that factors such as EPS production, expression of proteases and siderophores, and the metabolism of certain carbohydrates contribute to the different observed behaviors (Kube et al., 2010). The non-pathogenic species *E. billingiae* has also been observed to present antagonistic behavior in the presence of phytopathogenic fungi in *Pinus radiata*, where it is suggested that this *Erwinia* species can be used as a biocontrol agent (Mesanza et al., 2016). Recently, an endophytic isolate from potato has exhibited PGP traits such as phosphate solubilization and siderophore production. *In vivo* tests on potato seedlings confirmed the non-pathogenic behavior of this *Erwinia endophytica* isolate (Ramírez-Bahena et al., 2016). The specimens of *H. portulacoides* that were analyzed in the present work did not present symptoms of disease, which may suggest that the *Erwinia* spp. obtained in the Illumina-based sequencing belong to non-pathogenic species (Ramírez-Bahena et al., 2016).

For the family *Kiloniellaceae*, the traits observed in the endophytic collection have been discussed above for the genus *Thalassospira* in the AG core endophytome. In addition to our work, this genus has not been characterized regarding its PGP abilities or enzymatic activities. It has, however, been highly associated with the capacity to degrade hydrocarbons (e.g., Zhao et al, 2010), which can be a useful trait in the setting of a salt marsh subjected to influence of anthropogenic activities.

Other relevant families in the BG core endophytome included *Erythrobacteraceae* (28 OTUs, attributed to *Citromicrobium*, *Lutibacterium*, *Erythrobacter* and unassigned genera; 16314 reads), *Hyphomicrobiaceae* (19 OTUs, attributed to *Devosia* and unassigned genera; 13959 reads), *Pseudomonadaceae* (3 OTUs, attributed to *Pseudomonas* and an unassigned genus; 11693 reads), *Piscirickettsiaceae* (5 OTUs, attributed to *Methylophaga* and unassigned genera; 11071 reads), *Saccharosporillaceae* (1 unassigned OTU; 10813 reads), *Sphingomonadaceae* (14 OTUs, attributed to *Novosphingobium*, *Sphingomonas* and unassigned genera; 7447 reads), *Cohaesibacteraceae* (2 *Cohaesibacter* OTUs, 7165 reads), and *Rhizobiaceae* (7 OTUs, attributed to *Agrobacterium* and unassigned genera; 4730 reads).

Isolates from the family *Erythrobacteraceae* (assigned to the genera *Altererythrobacter*, *Citromicrobium*, *Erythrobacter* and *Erythromicrobium*) obtained in Chapter 3 were overall positive for all tested enzymatic activities and PGP traits. *Pseudomonas* isolates presented amylolytic, proteolytic, lipolytic (all isolates), cellulolytic, xylanolytic and pectinolytic (only at pH 7.0) activities, and PGP traits such as phosphate solubilization (7 out of 8 isolates), IAA production (including three isolates producing over 100  $\mu\text{g mL}^{-1}$ ), and ACC deaminase activity (all isolates). Only one isolate was assigned to the *Saccharospirillaceae* family, which belonged to the genus *Saccharospirillum*, and presented proteolytic and xylanolytic activities, and produced IAA. Isolates belonging to the family *Sphingomonadaceae* were assigned to the genera *Novosphingobium*, *Parasphingopyxis*, *Stakelama* and *Sphingorabdus*, and presented amylolytic, lipolytic, cellulolytic, xylanolytic and pectinolytic (only at pH 7.0) activities, and PGP traits such as phosphate solubilization (one *Stakelama* isolate), IAA production (all isolates), siderophore production, ACC deaminase activity and presence of the *nifH* gene. *Cohaesibacter* isolates presented amylolytic, lipolytic and cellulolytic activities, and were able to solubilize phosphate, produce IAA, present ACC deaminase activity, and the *nifH* gene was present. Finally, isolates from the family *Rhizobiaceae* included the genera *Ensifer* and *Rhizobium*, which presented amylolytic, proteolytic and pectinolytic activities (only *Ensifer* isolates), and lipolytic, cellulolytic and xylanolytic activities. PGP activities for this family included phosphate solubilization, siderophore and IAA production (one *Ensifer* isolate producing over 100  $\mu\text{g mL}^{-1}$ ), ACC deaminase activity and presence of the *nifH* gene.

In the culture-dependent methodology, no isolates were obtained from the families *Hyphomicrobiaceae* and *Piscirickettsiaceae* (Chapter 3). Members of the genus *Devosia* (family *Hyphomicrobiaceae*) have been isolated from plants from root nodules, are able to perform nitrate and nitrite reduction, and the genes for nodulation protein D and *nifH* have been detected, which are required for nodulation and symbiotic nitrogen fixation (Rivas et al., 2003). The genus *Methylophaga* (family *Piscirickettsiaceae*) has been associated with PGP traits such as IAA and ammonium production and ACC deaminase activity (Bal et al., 2013).

Interestingly, the AG core predominant family *Halomonadaceae* was not observed in the BG core, and the BG core families *Hyphomicrobiaceae*, *Saccharospirillaceae* and *Cohaesibacteraceae* were not observed in the AG core.

### **Does a rose, by any other name, smell just as sweet?**

The ever-growing application of high-throughput sequencing has triggered the need for automated bioinformatics tools that allow for the analysis of an extremely large amount of information in a small number of steps and in relatively little time. Two pipelines have dominated the field of bacterial 16S rRNA gene analysis obtained from high-throughput sequencing techniques: Mothur (Schloss et al., 2009) and QIIME (Caporaso et al., 2010). The latter was used in the present work, as is detailed in the Methods section of the present Chapter. There, it is also stated that the Greengenes database was used as the reference to obtain taxonomic affiliation of the OTUs that resulted from Illumina-based sequencing. Despite its outdated nature (last updated in 2013), this database is the default for the taxonomy attribution in the QIIME pipeline, and, for this reason, is commonly used when running the QIIME pipeline. The widespread use of the same database is beneficial in the sense that it allows for an easier and more reliable comparison of results, however, some attention should be given to the output it creates.

In the present work, one of our goals was to compare the diversity observed in the culture-based approach in Chapter 3 with the most abundant groups observed in a high-throughput sequencing approach. In doing so, several inaccuracies were observed in the attributions of the taxa names for several different OTUs, and in different taxonomic levels.

The taxonomic attribution of the lowest taxonomic level obtained from the Greengenes output was manually compared against the databases/platforms: LPSN (Euzéby, 1997), EzTaxon (Yoon et al., 2016), and SILVA (Quast et al., 2013). Overall, Greengenes attributed at least one inaccurate taxa level for approximately 20 % of our OTUs. Roughly half of these occurrences (393 OTUs) differ from what is consistently stated by the other databases -- e.g., in Greengenes, the genus *Marinilactibacillus* is attributed to the family *Aerococcaceae*, while the other databases file it under the family *Carnobacteriaceae*. In 315 OTUs the Greengenes attribution was also not consistent among the other databases -- e.g., the family *Mycobacteriaceae* is affiliated to the order Actinomycetales in the Greengenes and LPSN database, while it is under the order Corynebacteriales in the EzTaxon and SILVA databases. In the specific case of the genus *Anaerospira* (10 OTUs in our data), the misattribution occurred at all taxonomic levels (family, order, class and phylum) for this Firmicutes-affiliated genus that the Greengenes database had attributed to Proteobacteria.



Affiliation misattributions may have occurred for two main reasons: (i) in the four years since the Greengenes database was last updated, there may have been changes in taxa names -- e.g., species from the genus *Agrobacterium* (32 OTUs in our data) have been transferred to other genera, according to LPSN, SILVA and the International Committee on Systematics of Prokaryotes ([www.the-icsp.org](http://www.the-icsp.org)), and this genus no longer exists in the EzTaxon database; (ii) erroneous taxa names were introduced in the Greengenes database upon its construction, and have yet to be corrected. Alternatives to the Greengenes database should be sought after and applied, in order to have an updated and accurate record of taxonomic affiliations, nonetheless keeping in mind that no database appears to be exhaustively correct. The Society for General Microbiology ([www.microbiologyonline.org.uk](http://www.microbiologyonline.org.uk)) advocates for LPSN to find the correct name of a given bacterium, however this platform -- and/or its taxonomy browser -- is not wholly updated on a regular basis. Regardless of the used database, the output generated by these pipelines in terms of taxonomic affiliations should be critically examined before being discussed in publications.

## Conclusions

A considerably diverse community of bacteria resides in the inner tissues of *H. portulacoides*. Here, we observed a total of 3593 bacterial OTUs that comprised five main phyla and several minor ones, 85 classes, 182 families and 232 genera. We have determined the core bacterial endophytome of the halophyte using high-throughput sequencing and thorough analysis with high replication, and the results obtained evidence significant differences in the core endophytome in the AG and BG tissues.

Analysis of sequence similarity allowed us to identify several families in the core endophytome of AG and BG tissues in the present work. Previous works, including our isolation effort in Chapter 3, have exposed members of those families as pertaining several enzymatic activities and PGP traits, suggesting a significant ecological role for the diversity that is consistent in the halophyte.

The results obtained can be used to, not only comprehend the existing community, but also as a possible basis to predict effects of a disturbance, such as the presence of metals. The integrated knowledge obtained from trait analysis in the culture-dependent approach and the information retrieved from the culture-independent approach should be taken into account when considering the design of a bacterial consortium to apply in a PGP and/or phytoremediation setting.

## References

- Akinsanya, M. A., Goh, J. K., Lim, S. P. & Ting, A. S. Y. (2015). Metagenomics study of endophytic bacteria in *Aloe vera* using next-generation technology. *Genomics Data* **6**, 159-163.
- Amann, R. I., Ludwig, W. & Schleifer, K.-H. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* **59**, 143-169.
- Bal, H. B., Das, S., Dangar, T. K. & Adhya, T. K. (2013). ACC deaminase and IAA producing growth promoting bacteria from the rhizosphere soil of tropical rice plants. *J Basic Microbiol* **53**, 972-984.
- Bendich, A. J. (1987). Why do chloroplasts and mitochondria contain so many copies of their genome? *BioEssays* **6**, 279-282.
- Bodenhausen, N., Horton, M. W. & Bergelson, J. (2013). Bacterial communities associated with the leaves and roots of *Arabidopsis thaliana*. *PLoS One* **8**, e56329.
- Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L. & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* **64**, 807-838.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-336.
- Carrell, A. A., Carper, D. L. & Frank, A. C. (2016). Subalpine conifers in different geographical locations host highly similar foliar bacterial endophyte communities. *FEMS Microbiol Ecol* **92**, doi: 10.1093/femsec/fiw124.
- Chelius, M. K. & Triplett, E. W. (2001). The diversity of Archaea and Bacteria in association with the roots of *Zea mays* L.. *Microb Ecol* **41**, 252-263.
- Chen, C., Xin, K., Liu, H., Cheng, J., Shen, X., Wang, Y. & Zhang, L. (2017). *Pantoea alhagi*, a novel endophytic bacterium with ability to improve growth and drought tolerance in wheat. *Sci Rep* **7**, article 41564.
- de Abreu, C. S., Figueiredo, J. E. F., Oliveira, C. A., dos Santos, V. L., Gomes, E. A., Ribeiro, V. P., Barros, B. A., Lana, U. G. P. & Marriel, I. E. (2017). Maize endophytic bacteria as mineral phosphate solubilizers. *Genet Mol Res* **16**, doi:10.4238/gmr16019294.
- Delmont, T. O., Eren, A. M., Vineis, J. H. & Post, A. F. (2015). Genome reconstructions indicate the partitioning of ecological functions inside a phytoplankton bloom in the Amundsen Sea, Antarctica. *Front Microbiol* **6**, article 1090.

- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P. & Andersen, G. L. (2006).** Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb* **72**, 5069-5072.
- Edgar, R. C. (2010).** Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461.
- Ellis, R. J., Morgan, P., Weightman, A. J. & Fry, J. C. (2003).** Cultivation-dependent and -independent approaches for determining bacterial diversity in heavy-metal-contaminated soil. *Appl Environ Microb* **69**, 3223–3230.
- Euzéby, J. P. (1997).** List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet. *Int J Syst Bacteriol* **47**, 590-592.
- Fidalgo, C., Henriques, I., Rocha, J., Tacão, M. & Alves, A. (2016).** Culturable endophytic bacteria from the salt marsh plant *Halimione portulacoides*: phylogenetic diversity, functional characterization and influence of metal(loid) contamination. *Environ Sci Pollut Res* **23**:10200–10214.
- Glick, B. R. (2015).** Beneficial plant-bacterial interactions. Springer, Heidelberg.
- Gontia-Mishra, I., Sapre, S., Kachare, S. & Tiwari, S. (2017).** Molecular diversity of 1-aminocyclopropane-1-carboxylate (ACC) deaminase producing PGPR from wheat (*Triticum aestivum* L.) rhizosphere. *Plant Soil* **414**, 213-227.
- Gupta, A., Gopal, M., Thomas, G. V., Manikandan, V., Gajewski, J., Thomas, G., Seshagiri, S., Schuster, S. C., Rajesh, P. & Gupta, R. (2014).** Whole genome sequencing and analysis of plant growth promoting bacteria isolated from the rhizosphere of plantation crops Coconut, Cocoa and Arecanut. *PLoS One* **9**, article e104259.
- Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., Döring, M. & Sessitsch, A. (2015).** The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol Mol Biol Rev* **79**, 293-320.
- Hong, Y., Liao, D., Hu, A., Wang, H., Chen, J., Khan, S., Su, J. & Li, H. (2016).** Diversity of endophytic and rhizoplane bacterial communities associated with exotic *Spartina alterniflora* and native mangrove using Illumina amplicon sequencing. *Can J Microbiol* **61**, 723-733.
- Idris, R., Trifonova, R., Puschenreiter, M., Wenzel, W. W. & Sessitsch, A. (2004).** Bacterial communities associated with the flowering plants of the Ni hyperaccumulator *Thlaspi goesingense*. *Appl Environ Microbiol* **70**, 2667-2677.
- Jiao, J.-Y., Wang, H.-X., Zeng, Y. & Shen, Y.-M. (2006).** Enrichment for microbes living in association with plant tissues. *J Appl Microbiol* **100**, 830-837.

- Kube, M., Migdoll, A. M., Gehring, I., Heitmann, K., Mayer, Y., Kuhl, H., Knaust, F., Geider, K. & Reinhardt, R. (2010). Genome comparison of the epiphytic bacteria *Erwinia billingiae* and *E. tasmaniensis* with the pear pathogen *E. pyrifoliae*. *BMC Genomics* **11**, article 393.
- Lage, O. M. & Bondoso, J. (2011). Planctomycetes diversity associated with macroalgae. *FEMS Microbiol Ecol* **78**, 366-375.
- Lage, O. M. & Bondoso, J. (2012). Bringing Planctomycetes into pure culture. *Front Microbiol* **3**, 405.
- Lodewyckx, C., Vangronsveld, J., Porteous, F., Moore, E. R. B., Taghavi, S., Mezgeay, M. & van der Lelie, D. (2002). Endophytic bacteria and their potential applications. *Crit Rev Plant Sci* **21**, 583–606.
- Lucena, T., Mesa, J., Rodríguez-Llorente, I. D., Pajuelo, E., Caviedes, M. Á., Ruvira, M. A. & Pujalte, M. J. (2016). *Marinomonas spartinae* sp. nov., a novel species with plant-beneficial properties. *Int J Syst Evol Microbiol* **66**, 1686-1691.
- Lundberg, D. S., Yourstone, S., Mieczkowski, P., Jones, C. D. & Dangl, J. L. (2013). Practical innovations for high-throughput amplicon sequencing. *Nat Methods* **10**, 999-1002.
- Mapelli, F., Marasco, R., Rolli, E., Barbato, M., Cherif, H., Guesmi, A., Ouzari, I., Daffonchio, D. & Borin, S. (2013). Potential for plant growth promotion of rhizobacteria associated with *Salicornia* growing in Tunisian hypersaline soils. *Biomed Res Int* **2013**, 248078.
- Marcon, E. & Hérault, B. (2015). entropart: an {R} package to measure and partition diversity. *J Stat Softw* **67**, 1-26.
- McMurdie, P. J. & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217.
- McMurdie, P. J. & Holmes, S. (2014). Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* **10**, e1003531.
- Mesa, J., Mateos-Naranjo, E., Caviedes, M. A., Redondo-Gómez, S., Pajuelo, E. & Rodríguez-Llorente, I. D. (2015). Endophytic cultivable bacteria of the metal bioaccumulator *Spartina maritima* improve plant growth but not metal uptake in polluted marshes soils. *Front Microbiol* **6**, article 1450.
- Mesanza, N., Iturrutxa, E. & Patten, C. L. (2016). Native rhizobacteria as biocontrol agents of *Heterobasidion annosum* s.s. and *Armillaria mellea* infection of *Pinus radiata*. *Biol Control* **101**, 8-16.
- Mora-Ruiz, M. del R., Font-Verdera, F., Díaz-Gil, C., Urdiain, M., Rodríguez-Valdecantos, G., González, B., Orfila, A. & Rosselló-Móra, R. (2015). Moderate halophilic bacteria colonizing the phylloplane of halophytes of the subfamily *Salicornioideae* (Amaranthaceae). *Syst Appl Microbiol* **38**, 406-416.

- Mora-Ruiz, M. del R., Font-Verdera, F., Orfila, A., Rita, J. & Rosselló-Móra, R. (2016).** Endophytic microbial diversity of the halophyte *Arthrocnemum macrostachyum* across plant compartments. *FEMS Microbiol Ecol* **92**, doi: 10.1093/femsec/fiw145.
- Musat, N., Werner, U., Knittel, K., Kolb, S., Dodenhof, T., van Beusekom, J. E., de Beer, D., Dubilier, N. & Amann, R. (2006).** Microbial community structure of sandy intertidal sediments in the North Sea, SyltRømø Basin, Wadden Sea. *Syst Appl Microbiol* **29**, 333-348.
- Muyzer, G., de Waal, E. C. & Uitterlinden, A. G. (1993).** Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**, 695-700.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E. & Wagner, H. (2016).** Vegan: community ecology package. <https://CRAN.R-project.org/package=vegan>
- Oliveira, V., Gomes, N. C. M., Almeida, A., Silva, A. M. S., Simões, M. Q. M., Smalla, K. & Cunha, A. (2014a).** Hydrocarbon contamination and plant species determine the phylogenetic and functional diversity of endophytic degrading bacteria. *Mol Ecol* **23**, 1392-1404.
- Oliveira, V., Gomes, N. C. M., Cleary, D. F. R., Almeida, A., Silva, A. M. S., Simões, M. M. Q., Silva, H. & Cunha, A. (2014b).** Halophyte plant colonization as a driver of the composition of bacterial communities in salt marshes chronically exposed to oil hydrocarbons. *FEMS Microbiol Ecol* **90**, 647-662.
- Philippot, L., Raaijmakers, J. M., Lemanceau, P. & van der Putten, W. H. (2013).** Going back to the roots: the microbial ecology of the rhizosphere. *Nat Rev Microbiol* **11**, 789-799.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glöckner, F. O. (2013).** The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl Acids Res* **41**, D590-D596.
- R Core Team (2016).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Ramírez-Bahena, M. H., Salazar, S., Cuesta, M. J., Tejedor, C., Igual, J.-M., Fernández-Pascual, M. & Peix, A. (2016).** *Erwinia endophytica* sp. nov., isolated from potato (*Solanum tuberosum* L.) stems. *Int J Syst Evol Microbiol* **66**, 975-981.
- Ribeiro, C. M. & Cardoso, E. J. B. N. (2012).** Isolation, selection and characterization of root-associated growth promoting bacteria in Brazil Pine (*Araucaria angustifolia*). *Microbiol Res* **167**, 69-78.
- Rideout, J. R., He, Y., Navas-Molina, J. A., Walters, W. A., Ursell, L. K., Gibbons, S. M., Chase, J., McDonald, D., Gonzalez, A., Robbins-Pianka, A., Clemente, J. C., Gilbert, J. A., Huse, S. M., Zhou, H.-W., Knight, R. &**

- Caporaso, J. G. (2014).** Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* **2**, e545; doi:10.7717/peerj.545.
- Rivas, R., Willems, A., Subba-Rao, N. S., Mateos, P. F., Dazzo, F. B., Kroppenstedt, R. M., Martínez-Molina, E., Gillis, M. & Velázquez, E. (2003).** Description of *Devosia neptuniae* sp. nov. that nodulates and fixes nitrogen in symbiosis with *Neptunia natans*, an aquatic legume from India. *Appl Environ Microbiol* **26**, 47-53.
- Ruiz-Pérez, C. A., Restrepo, S. & Zambrano, M. M. (2016).** Microbial and Functional Diversity within the Phyllosphere of Espeletia sp. in an Andean High Mountain Ecosystem. *Appl Environ Microbiol* **82**, 1807-1817.
- Santoyo, G., Moreno-Hagelsieb, G., del Carmen Orozco-Mosqueda, M. & Glick, B. R. (2016).** Plant growth-promoting bacterial endophytes. *Microbiol Res* **183**, 92-99.
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. & Weber, C. F. (2009).** Introducing Mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, 7537-7541.
- Schmieder, R. & Edwards, R. (2011).** Quality control and preprocessing of metagenomics datasets. *Bioinformatics* **27**, 863-864.
- Schubert M, Lindgreen S, Orlando L. 2016.** AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* **9**, 88.
- Shade, A. & Handelsman, J. (2011).** Beyond the Venn diagram: the hunt for a core microbiome. *Environ Microbiol* **14**, 4-12.
- Shen, S. Y. & Fulthorpe, R. (2015).** Seasonal variation of bacterial endophytes in urban trees. *Front Microbiol* **6**, 427. doi: 10.3389/fmicb.2015.00427
- Shuang., J. L., Zhang, X. Y., Zhao, Z. Z., Yao, S. P., An, S. Q., Xue, Y. R. & Liu, C. H. (2009).** Bacterial phylogenetic diversity in a Spartina marsh in China. *Ecol Eng* **35**, 529-535.
- Skirnisdottir, S., Hreggvidsson, G. O., Hjörleifsdottir, S., Marteinson, V. T., Petursdottir, S. K., Holst, O. & Kristjansson, J. K. (2000).** Influence of sulfide and temperature on species composition and community structure of hot spring microbial mats. *Appl Environ Microbiol* **66**, 2835-2841.
- Su, J., Ouyang, W., Hong, Y., Liao, D., Khan, S. & Li, H. (2016).** Responses of endophytic and rhizospheric bacterial communities of salt marsh plant (*Spartina alterniflora*) to polycyclic aromatic hydrocarbons contamination. *J Soils Sediments* **16**, 707-715.

- Szymańska, S., Płociniczak, T., Piotrowska-Seget, Z. & Hryniewicz, K. (2016a).** Endophytic and rhizosphere bacteria associated with the roots of the halophyte *Salicornia europaea* L. - community structure and metabolic potential. *Microbiol Res* **192**, 37-51.
- Szymańska, S., Płociniczak, T., Piotrowska-Seget, Z., Złoch, M., Ruppel, S. & Hryniewicz, K. (2016b).** Metabolic potential and community structure of endophytic and rhizosphere bacteria associated with the roots of the halophyte *Aster tripolium* L. *Microbiol Res* **182**, 68-79.
- Tanaka, T., Kawasaki, K., Daimon, S., Kitagawa, W., Yamamoto, K., Tamaki, H., Tanaka, M., Nakatsu, C. H. & Kamagata, Y. (2014).** A hidden pitfall in agar media preparation undermines cultivability of microorganisms. *Appl Environ Microbiol* **80**, 7659–7666.
- Thomas, P. & Upreti, R. (2016).** Evaluation of tomato seedling root-associated bacterial endophytes towards organic seedling production. *Org Agr* **6**, 89-98.
- Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. (2015).** The importance of the microbiome of the plant holobiont. *New Phytol* **206**, 1196-1206.
- Wang, H. X., Geng, Z. L., Zeng, Y. & Shen, Y. M. (2008).** Enriching plant microbiota for a metagenomic library construction. *Environ Microbiol* **10**, 2684-2691.
- Wang, W., Zhai, Y., Cao, L., Tan, H. & Zhang, R. (2016a).** Endophytic bacterial and fungal microbiota in sprouts, roots and stems of rice (*Oryza sativa* L.). *Microbiol Res* **188-189**, 1-8.
- Wang, Y., Ling, X., Gu, Y. Q. & Coleman-Derr, D. (2016b).** MetaCoMET: a web platform for discovery and visualization of the core microbiome. *Bioinformatics* **32**, 3469-3470.
- Xiao, X., Chen, W., Zong, L., Yang, J., Jiao, S., Lin, Y., Wang, E. & Wei, G. (2017).** Two cultivated legume plants reveal the enrichment process of the microbiome in the rhizocompartments. *Mol Ecol* **26**, 1641-1651.
- Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J. (2016).** Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. *Int J Syst Evol Microbiol*; doi: 10.1099/ijsem.0.001755
- Zhao, B., Wang, H., Li, R. & Mao, X. (2010).** *Thalassospira xianhensis* sp. nov., a polycyclic aromatic hydrocarbon-degrading marine bacterium. *Int J Syst Evol Microbiol* **60**, 1125-1129.
- Zhao, S., Zhou, N., Zhao, Z.-Y., Zhang, K. & Tian, C.-Y. (2016).** High-throughput sequencing analysis of the endophytic bacterial diversity and dynamics in roots of the halophyte *Salicornia europaea*. *Curr Microbiol* **72**, 557-562.



## Annex I

### Function *pairwise.adonis*

```
pairwise.adonis <- function(x, factors, sim.method = 'bray', p.adjust.m = 'bonferroni')
{
  library(vegan)
  co = combn(unique(factors),2)
  pairs = c()
  F.Model = c()
  R2 = c()
  p.value = c()

  for(elem in 1:ncol(co)){
    ad = adonis(x[factors %in% c(co[1,elem],co[2,elem]),] ~ factors[factors %in%
c(co[1,elem],co[2,elem])], method = sim.method);
    pairs = c(pairs,paste(co[1,elem],'vs',co[2,elem]));
    F.Model = c(F.Model,ad$aoi.tab[1,4]);
    R2 = c(R2,ad$aoi.tab[1,5]);
    p.value = c(p.value,ad$aoi.tab[1,6])
  }
  p.adjusted = p.adjust(p.value,method=p.adjust.m)
  pairw.res = data.frame(pairs,F.Model,R2,p.value,p.adjusted)
  return(pairw.res)
}
```

## Annex II

**Table 5.II.1** List of phyla found in the endophytic bacterial community of *H. portulacoides*.

Phylum	OTUs in phylum	Percentage of OTUs in phylum
[Caldithrix]	1	0.03%
[Thermi]	4	0.11%
Acidobacteria	64	1.78%
Actinobacteria	185	5.15%
Aquificae	2	0.06%
Armatimonadetes	1	0.03%
Bacteroidetes	385	10.72%
BRC1	4	0.11%
Chlamydiae	12	0.33%
Chlorobi	5	0.14%
Chloroflexi	39	1.09%
Cyanobacteria	17	0.47%
Elusimicrobia	1	0.03%
FBP	1	0.03%
Fibrobacteres	5	0.14%
Firmicutes	85	2.37%
Fusobacteria	3	0.08%
Gemmatimonadetes	15	0.42%
GN02	8	0.22%
Kazan-3B-28	1	0.03%
Lentisphaerae	3	0.08%
Nitrospirae	3	0.08%
NKB19	8	0.22%
OD1	14	0.39%
OP11	9	0.25%
Planctomycetes	316	8.79%
Proteobacteria	2244	62.45%
SBR1093	2	0.06%
Spirochaetes	13	0.36%
Tenericutes	3	0.08%
TM6	4	0.11%
TM7	21	0.58%
Verrucomicrobia	103	2.87%
WPS-2	5	0.14%
WS2	3	0.08%
WS3	3	0.08%
WS6	1	0.03%

**Table 5.II.2** List of classes found in the endophytic bacterial community of *H. portulacoides*.

Class	OTUs in class	Percentage of OTUs in class
n.a.	20	0.56%
[Brachyspirae]	1	0.03%
[Chloracidobacteria]	10	0.28%
[Lentisphaeria]	3	0.08%
[Leptospirae]	5	0.14%
[Methyloacidiphilae]	1	0.03%
[Pedosphaerae]	7	0.19%
[Rhodothermi]	14	0.39%
[Saprospirae]	48	1.34%
[Spartobacteria]	11	0.31%
0319-6E2	1	0.03%

3BR-5F	2	0.06%
4C0d-2	4	0.11%
ABY1	4	0.11%
Acidimicrobiia	66	1.84%
Acidobacteria-6	12	0.33%
Acidobacteriia	2	0.06%
Actinobacteria	98	2.73%
Alphaproteobacteria	1036	28.83%
Anaerolineae	15	0.42%
Aquificae	2	0.06%
AT-s54	1	0.03%
B142	1	0.03%
Bacilli	48	1.34%
Bacteroidia	12	0.33%
BB34	2	0.06%
BD7-11	2	0.06%
Betaproteobacteria	50	1.39%
BME43	1	0.03%
C6	4	0.11%
Chlamydiia	12	0.33%
Chloroflexi	1	0.03%
Clostridia	37	1.03%
Cytophagia	111	3.09%
Deinococci	4	0.11%
Deltaproteobacteria	132	3.67%
Ellin6529	5	0.14%
Elusimicrobia	1	0.03%
Epsilonproteobacteria	11	0.31%
Fibrobacteria	5	0.14%
Flavobacteriia	192	5.34%
Fusobacteriia	3	0.08%
Gammaproteobacteria	1015	28.25%
Gemm-1	1	0.03%
Gemm-2	12	0.33%
Gemm-3	1	0.03%
GKS2-174	3	0.08%
GN07	1	0.03%
Holophagae	5	0.14%
KSB1	1	0.03%
Ktedonobacteria	1	0.03%
ML635J-21	2	0.06%
Mollicutes	3	0.08%
Nitriliruptoria	12	0.33%
Nitrospira	3	0.08%
NPL-UPA2	1	0.03%
OM190	16	0.45%
OPB56	2	0.06%
Opitutae	33	0.92%
Oscillatoriothycideae	8	0.22%
Phycisphaerae	29	0.81%
Planctomycetia	259	7.21%
PRR-11	3	0.08%
PRR-12	3	0.08%
RB25	1	0.03%
Rubrobacteria	1	0.03%
S085	6	0.17%
SHA-109	3	0.08%
SJA-28	3	0.08%
SJA-4	4	0.11%
Solibacteres	6	0.17%
Sphingobacteriia	7	0.19%

Spirochaetes	7	0.19%
Sva0725	26	0.72%
Synechococcophycideae	3	0.08%
Thermoleophilia	8	0.22%
Thermomicrobia	3	0.08%
TK17	8	0.22%
TM7-1	16	0.45%
TM7-3	2	0.06%
TSBW08	4	0.11%
vadinHA49	2	0.06%
Verrucomicrobiae	51	1.42%
VHS-B5-50	2	0.06%
WCHB1-64	9	0.25%
ZB2	10	0.28%

n.a., not attributed.

**Table 5.II.3** List of orders found in the endophytic bacterial community of *H. portulacoides*.

Order	OTUs in order	Percentage of OTUs in order
n.a.	219	6.12%
[Brachyspirales]	1	0.03%
[Cerasicoccales]	5	0.14%
[Chthoniobacterales]	11	0.31%
[Leptospirales]	5	0.14%
[Marinicellales]	10	0.28%
[Pedosphaerales]	4	0.11%
[Pelagicoccales]	5	0.14%
[Rhodothermales]	14	0.39%
[Saprospirales]	48	1.34%
258ds10	5	0.14%
34P16	4	0.11%
Acidimicrobiales	66	1.85%
Acidobacteriales	2	0.06%
Actinomycetales	97	2.71%
Aeromonadales	2	0.06%
agg27	9	0.25%
AKIW781	1	0.03%
AKYG1722	2	0.06%
Alteromonadales	229	6.40%
Aquificales	2	0.06%
B97	3	0.08%
Bacillales	16	0.45%
Bacteroidales	12	0.34%
BD7-3	39	1.09%
Bdellovibrionales	35	0.98%
Bifidobacteriales	1	0.03%
Burkholderiales	33	0.92%
Caldilineales	5	0.14%
Campylobacteriales	11	0.31%
Caulobacteriales	8	0.22%
CCM11a	3	0.08%
CCU21	3	0.08%
Chlamydiales	12	0.34%
Chromatiales	14	0.39%
Chroococcales	8	0.22%
CL500-15	7	0.20%
Clostridiales	37	1.03%
Cytophagales	111	3.10%

d113	3	0.08%
d153	8	0.22%
Deinococcales	4	0.11%
Desulfobacterales	2	0.06%
Desulfovibrionales	4	0.11%
Ellin329	1	0.03%
Enterobacteriales	205	5.73%
Entomoplasmatales	1	0.03%
Euzebyales	5	0.14%
EW055	1	0.03%
Flavobacteriales	192	5.37%
Fusobacteriales	3	0.08%
Gaiellales	2	0.06%
Gemmatales	5	0.14%
GMD14H09	1	0.03%
HOC36	9	0.25%
Holophagales	5	0.14%
HTCC2188	6	0.17%
Hydrogenophilales	3	0.08%
Ilb	1	0.03%
iii1-15	9	0.25%
Kiloniellales	56	1.57%
Kordiimonadales	2	0.06%
Ktedonobacteriales	1	0.03%
Lactobacillales	32	0.89%
Legionellales	36	1.01%
Lentisphaerales	3	0.08%
Methylococcoides	1	0.03%
Methylophilales	3	0.08%
MLE1-12	3	0.08%
mle1-48	2	0.06%
MVS-107	1	0.03%
Mycoplasmatales	2	0.06%
Myxococcales	50	1.40%
NB1-j	14	0.39%
Neisseriales	1	0.03%
Nitriliruptorales	7	0.20%
Nitrosomonadales	2	0.06%
Nitrospirales	3	0.08%
Oceanospirillales	268	7.49%
Opitutales	9	0.25%
PB19	4	0.11%
Phycisphaerales	23	0.64%
Pirellulales	157	4.39%
Planctomycetales	94	2.63%
Pseudanabaenales	3	0.08%
Pseudomonadales	87	2.43%
Puniceococcales	6	0.17%
RB41	10	0.28%
Rhizobiales	284	7.94%
Rhodobacterales	218	6.09%
Rhodocyclales	6	0.17%
Rhodospirillales	69	1.93%
Rickettsiales	43	1.20%
Rubrobacteriales	1	0.03%
S0208	2	0.06%
S-70	1	0.03%
Salinisphaerales	20	0.56%
SBR1031	7	0.20%
Sediment-1	3	0.08%
SM1D11	1	0.03%

Solibacterales	6	0.17%
Solirubrobacterales	6	0.17%
Sphingobacterales	7	0.20%
Sphingomonadales	255	7.13%
Spirobacillales	12	0.34%
Spirochaetales	6	0.17%
Sva0725	26	0.73%
Sva0853	1	0.03%
Syntrophobacterales	3	0.08%
Thiohalorhabdales	5	0.14%
Thiotrichales	47	1.31%
Ucn15732	1	0.03%
Verrucomicrobiales	51	1.43%
Vibrionales	24	0.67%
WD2101	2	0.06%
Xanthomonadales	21	0.59%

n.a., not attributed.

**Table 5.II.4** List of families found in the endophytic bacterial community of *H. portulacoides*.

Family	OTUs in family	Percentage of OTUs in family
n.a.	660	18.64%
[Amoebophilaceae]	11	0.31%
[Balneolaceae]	3	0.08%
[Cerasioccaceae]	5	0.14%
[Chromatiaceae]	6	0.17%
[Chthoniobacteraceae]	11	0.31%
[Marinicellaceae]	10	0.28%
[Pelagicoccaceae]	5	0.14%
[Thermodesulfovibrionaceae]	2	0.06%
[Tissierellaceae]	23	0.65%
[Weeksellaceae]	8	0.23%
211ds20	14	0.40%
A4b	5	0.14%
Acetobacteraceae	2	0.06%
Acidobacteriaceae	1	0.03%
Actinomycetaceae	1	0.03%
Aerococcaceae	17	0.48%
Aeromonadaceae	2	0.06%
Alcaligenaceae	2	0.06%
Alcanivoracaceae	3	0.08%
Alicyclobacillaceae	3	0.08%
Alteromonadaceae	147	4.15%
Aquificaceae	1	0.03%
Arctic95B-10	3	0.08%
Aurantimonadaceae	10	0.28%
Bacillaceae	3	0.08%
Bacteriovoracaceae	21	0.59%
Bacteroidaceae	4	0.11%
Bdellovibrionaceae	14	0.40%
Beijerinckiaceae	1	0.03%
Bifidobacteriaceae	1	0.03%
Brachyspiraceae	1	0.03%
Bradyrhizobiaceae	4	0.11%
Brevibacteriaceae	1	0.03%
Brucellaceae	1	0.03%
C111	14	0.40%
Caldilineaceae	5	0.14%

Campylobacteraceae	11	0.31%
Carnobacteriaceae	1	0.03%
Caulobacteraceae	8	0.23%
Cellulomonadaceae	9	0.25%
Chitinophagaceae	22	0.62%
Clostridiaceae	1	0.03%
Cohaesibacteraceae	20	0.56%
Colwelliaceae	2	0.06%
Comamonadaceae	21	0.59%
Corynebacteriaceae	8	0.23%
Coxiellaceae	12	0.34%
Criblamydiaceae	1	0.03%
Cryomorphaceae	19	0.54%
CV106	1	0.03%
Cyclobacteriaceae	5	0.14%
Cystobacterineae	1	0.03%
Cytophagaceae	20	0.56%
Dermabacteraceae	2	0.06%
Dermacoccaceae	2	0.06%
Desulfobulbaceae	1	0.03%
Desulfovibrionaceae	4	0.11%
Ectothiorhodospiraceae	1	0.03%
Ellin6075	9	0.25%
Enterobacteriaceae	205	5.79%
Enterococcaceae	2	0.06%
Erythrobacteraceae	124	3.50%
Euzebyaceae	5	0.14%
Ferrimonadaceae	1	0.03%
Flammeovirgaceae	74	2.09%
Flavobacteriaceae	155	4.38%
Francisellaceae	1	0.03%
Fusobacteriaceae	2	0.06%
Gemmataceae	3	0.08%
Geodermatophilaceae	3	0.08%
Gordoniaceae	1	0.03%
Hahellaceae	4	0.11%
Haliangiaceae	2	0.06%
Halomonadaceae	48	1.36%
Halothiobacillaceae	1	0.03%
Holosporaceae	1	0.03%
HTCC2089	2	0.06%
HTCC2188	14	0.40%
Hydrogenophilaceae	3	0.08%
Hydrogenothermaceae	1	0.03%
Hyphomicrobiaceae	114	3.22%
Hyphomonadaceae	12	0.34%
Idiomarinaceae	8	0.23%
Intrasporangiaceae	8	0.23%
Isosphaeraceae	2	0.06%
JdFBGBact	1	0.03%
Jonesiaceae	1	0.03%
JTB36	1	0.03%
JTB38	6	0.17%
Kiloniellaceae	45	1.27%
Kineosporiaceae	2	0.06%
koll13	3	0.08%
Kordiimonadaceae	2	0.06%
Koribacteraceae	1	0.03%
Ktedonobacteraceae	1	0.03%
Lachnospiraceae	5	0.14%
Lactobacillaceae	1	0.03%

Legionellaceae	10	0.28%
Leptospiraceae	5	0.14%
Leptotrichiaceae	1	0.03%
Leuconostocaceae	4	0.11%
Methylobacteriaceae	7	0.20%
Methylocystaceae	2	0.06%
Methylophilaceae	3	0.08%
Microbacteriaceae	6	0.17%
Micrococcaceae	4	0.11%
Microthrixaceae	4	0.11%
MND4	1	0.03%
Moraxellaceae	20	0.56%
Mycobacteriaceae	9	0.25%
Mycoplasmataceae	2	0.06%
Nannocystaceae	4	0.11%
NB1-i	3	0.08%
Neisseriaceae	1	0.03%
Nitriliruptoraceae	7	0.20%
Nitrosomonadaceae	2	0.06%
Nitrospinaceae	1	0.03%
Nitrospiraceae	1	0.03%
Nocardiaceae	3	0.08%
Nocardioidaceae	15	0.42%
ntu14	1	0.03%
Oceanospirillaceae	192	5.42%
Oleiphilaceae	2	0.06%
OM60	11	0.31%
Opitutaceae	9	0.25%
Oxalobacteraceae	10	0.28%
Paenibacillaceae	3	0.08%
Parachlamydiaceae	2	0.06%
PAUC26f	5	0.14%
Phycisphaeraceae	9	0.25%
Phyllobacteriaceae	25	0.71%
Pirellulaceae	157	4.43%
Piscirickettsiaceae	47	1.33%
Planctomycetaceae	94	2.65%
Planococcaceae	2	0.06%
Polyangiaceae	1	0.03%
Porphyromonadaceae	3	0.08%
Prevotellaceae	3	0.08%
Promicromonosporaceae	1	0.03%
Propionibacteriaceae	6	0.17%
Pseudanabaenaceae	3	0.08%
Pseudoalteromonadaceae	3	0.08%
Pseudomonadaceae	66	1.86%
Pseudonocardiaceae	8	0.23%
Puniceicoccaceae	6	0.17%
RB40	6	0.17%
Rhabdochlamydiaceae	1	0.03%
Rhizobiaceae	56	1.58%
Rhodobacteraceae	206	5.82%
Rhodobiaceae	2	0.06%
Rhodocyclaceae	6	0.17%
Rhodospirillaceae	51	1.44%
Rhodothermaceae	11	0.31%
Rickettsiaceae	27	0.76%
Rubrobacteraceae	1	0.03%
Ruminococcaceae	4	0.11%
Saccharospirillaceae	17	0.48%
Salinisphaeraceae	20	0.56%



Sanguibacteraceae	1	0.03%
Saprospiraceae	25	0.71%
SC3-41	3	0.08%
Shewanellaceae	5	0.14%
Simkaniaceae	2	0.06%
Sinobacteraceae	5	0.14%
SJA-101	2	0.06%
Solirubrobacteraceae	1	0.03%
Sphingobacteriaceae	3	0.08%
Sphingomonadaceae	100	2.82%
Spirochaetaceae	6	0.17%
Sporichthyaceae	1	0.03%
Staphylococcaceae	4	0.11%
Streptococcaceae	6	0.17%
Syntrophobacteraceae	3	0.08%
Thermomonosporaceae	1	0.03%
TK06	2	0.06%
Trueperaceae	4	0.11%
Veillonellaceae	4	0.11%
Verrucomicrobiaceae	51	1.44%
Vibrionaceae	21	0.59%
wb1_P06	4	0.11%
Xanthomonadaceae	16	0.45%
Xenococcaceae	8	0.23%

n.a., not attributed.

**Table 5.II.5** List of genera found in the endophytic bacterial community of *H. portulacoides*.

Genera	OTUs in genus	Percentage of OTUs in genus
n.a.	1842	54.66%
1-68	1	0.03%
A17	9	0.27%
Acinetobacter	10	0.30%
Actinomyces	1	0.03%
Actinomycetospora	1	0.03%
Actinotalea	4	0.12%
Aequorivita	4	0.12%
Aeromicrobium	2	0.06%
Afifella	1	0.03%
Agrobacterium	32	0.95%
Alcanivorax	3	0.09%
Alicyclobacillus	3	0.09%
Alloioicoccus	1	0.03%
Alteromonas	4	0.12%
Amaricoccus	1	0.03%
Amphritea	2	0.06%
Anaerococcus	10	0.30%
Anaerospira	10	0.30%
Antarctobacter	1	0.03%
Aquicella	3	0.09%
Aquimonas	1	0.03%
Arcobacter	8	0.24%
Arenibacter	1	0.03%
Aurantimonas	1	0.03%
B-42	2	0.06%
Bacillus	2	0.06%
Bacteriovorax	7	0.21%
Bacteroides	4	0.12%

Balneola	3	0.09%
BD2-13	33	0.98%
Bdellovibrio	14	0.42%
Bifidobacterium	1	0.03%
Blautia	1	0.03%
Brachybacterium	1	0.03%
Brevibacterium	1	0.03%
Brumimicrobium	2	0.06%
Campylobacter	2	0.06%
Candidatus Cardinium	1	0.03%
Candidatus Endobugula	1	0.03%
Candidatus Portiera	3	0.09%
Candidatus Rhabdochlamydia	1	0.03%
Candidatus Xiphinematobacter	4	0.12%
Capnocytophaga	1	0.03%
Carnobacterium	1	0.03%
Cellvibrio	19	0.56%
Cerasicoccus	1	0.03%
Chryseobacterium	4	0.12%
Chthoniobacter	2	0.06%
Citrobacter	155	4.60%
Citromicrobium	10	0.30%
Cloacibacterium	2	0.06%
Clostridium	1	0.03%
Cohaesibacter	20	0.59%
Comamonas	1	0.03%
Coproccoccus	1	0.03%
Coralimargarita	3	0.09%
Corynebacterium	8	0.24%
Crocinitomix	5	0.15%
Cupriavidus	3	0.09%
DA101	1	0.03%
Dechloromonas	1	0.03%
Delftia	4	0.12%
Demequina	3	0.09%
Dermabacter	1	0.03%
Dermacoccus	2	0.06%
Desulfovibrio	4	0.12%
Devosia	58	1.72%
Dialister	1	0.03%
Dinoroseobacter	1	0.03%
Dokdonella	4	0.12%
Elizabethkingia	1	0.03%
Enhydrobacter	1	0.03%
Enterococcus	1	0.03%
Enterovibrio	1	0.03%
Erwinia	20	0.59%
Erythrobacter	6	0.18%
Euzebya	5	0.15%
Faecalibacterium	1	0.03%
Ferrimonas	1	0.03%
FFCH10602	1	0.03%
Finegoldia	4	0.12%
Flavobacterium	19	0.56%
Flexibacter	1	0.03%
Flexithrix	2	0.06%
Fluviicola	5	0.15%
Fulvivirga	12	0.36%
Fusobacterium	1	0.03%
Gemmata	1	0.03%
Glaciecola	18	0.53%

Gordonia	1	0.03%
Gramella	1	0.03%
Haererehalobacter	1	0.03%
Hahella	4	0.12%
Haloferula	2	0.06%
Halomonas	41	1.22%
HTCC	13	0.39%
Hydrogenophaga	2	0.06%
Hymenobacter	4	0.12%
Hyphomonas	2	0.06%
Idiomarina	8	0.24%
Inquilinus	2	0.06%
Jannaschia	2	0.06%
Jeotgalicoccus	1	0.03%
JTB248	2	0.06%
Kaistobacter	1	0.03%
Kineococcus	2	0.06%
Klebsiella	2	0.06%
Kocuria	1	0.03%
Kordia	1	0.03%
Kushneria	1	0.03%
Lactobacillus	1	0.03%
Leadbetterella	1	0.03%
Leeuwenhoekiella	10	0.30%
Legionella	2	0.06%
Leptonema	5	0.15%
Leptotrichia	1	0.03%
Leuconostoc	4	0.12%
Lewinella	10	0.30%
Loktanella	1	0.03%
Luteimonas	2	0.06%
Luteolibacter	18	0.53%
Lutibacterium	9	0.27%
Lysobacter	1	0.03%
Maribacter	3	0.09%
Marinicella	6	0.18%
Marinilactibacillus	16	0.47%
Marinimicrobium	1	0.03%
Marinobacter	27	0.80%
Marinobacterium	1	0.03%
Marinomonas	168	4.99%
Marivita	1	0.03%
Megamonas	1	0.03%
Mesonina	5	0.15%
Mesorhizobium	1	0.03%
Methylibium	2	0.06%
Methylobacterium	5	0.15%
Methylophaga	17	0.50%
Methylotheria	3	0.09%
Microbacterium	4	0.12%
Microbulbifer	3	0.09%
Micrococcus	1	0.03%
Muricauda	12	0.36%
Mycobacterium	9	0.27%
Mycoplasma	2	0.06%
ND137	3	0.09%
Neptunomonas	1	0.03%
Nesterenkonia	1	0.03%
Nisaea	2	0.06%
Nitratireductor	1	0.03%
Nitrocola	2	0.06%

Nitrospina	1	0.03%
Nocardioides	1	0.03%
Novosphingobium	26	0.77%
Oceanicaulis	1	0.03%
Oceanospirillum	1	0.03%
Ochrobactrum	1	0.03%
Oleibacter	3	0.09%
Opitutus	5	0.15%
Owenweeksia	3	0.09%
Paenibacillus	2	0.06%
Paracoccus	9	0.27%
Parvibaculum	2	0.06%
Pelagicoccus	5	0.15%
Peptoniphilus	7	0.21%
Persicitalea	1	0.03%
ph2	1	0.03%
Phycococcus	1	0.03%
Phyllobacterium	2	0.06%
Pigmentiphaga	1	0.03%
Pimelobacter	1	0.03%
Planctomyces	94	2.79%
Planococcus	1	0.03%
Pleomorphomonas	1	0.03%
Plesiocystis	4	0.12%
Polaribacter	4	0.12%
Porphyromonas	3	0.09%
Prevotella	3	0.09%
Propionibacterium	4	0.12%
Pseudoalteromonas	3	0.09%
Pseudomonas	57	1.69%
Pseudonocardia	7	0.21%
Psychrilyobacter	1	0.03%
Psychrobacter	5	0.15%
Psychroserpens	1	0.03%
Puniceicoccus	3	0.09%
Ralstonia	3	0.09%
Reichenbachiella	2	0.06%
Rheinheimera	2	0.06%
Rhodobacter	1	0.03%
Rhodobium	1	0.03%
Rhodococcus	2	0.06%
Rhodoplanes	2	0.06%
Rhodospirillum	2	0.06%
Rhodovibrio	1	0.03%
Robiginitalea	1	0.03%
Roseburia	1	0.03%
Roseivirga	1	0.03%
Rubricoccus	9	0.27%
Rubrobacter	1	0.03%
Ruminococcus	2	0.06%
Saccharospirillum	2	0.06%
Salegentibacter	4	0.12%
Salinisphaera	20	0.59%
Sanguibacter	1	0.03%
SC3-56	9	0.27%
Sediminibacterium	1	0.03%
Shewanella	5	0.15%
Sphingobacterium	1	0.03%
Sphingobium	5	0.15%
Sphingomonas	8	0.24%
Sphingopyxis	1	0.03%

Spirochaeta	5	0.15%
Spirosoma	2	0.06%
Sporocytophaga	1	0.03%
Staphylococcus	3	0.09%
Stenotrophomonas	2	0.06%
Steroidobacter	3	0.09%
Streptococcus	6	0.18%
Sulfitobacter	3	0.09%
Tenacibaculum	4	0.12%
Tepidimonas	1	0.03%
Thalassomonas	2	0.06%
Thalassospira	44	1.31%
Thiobacillus	3	0.09%
Thiovirga	1	0.03%
Ulvibacter	1	0.03%
Veillonella	2	0.06%
Verrucomicrobium	6	0.18%
Vibrio	7	0.21%
Winogradskyella	3	0.09%
Wolbachia	3	0.09%
Xylanimicrobium	1	0.03%

n.a., not attributed.



## Chapter 6

---

Final considerations

## **Contents**

### **6.1** Final considerations

### **6.2** Conclusions

### **6.3** Future work



## Final considerations

### Considerations regarding metal(loid) contamination in Ria de Aveiro

The contamination observed in Ria de Aveiro, especially in the Laranjo Bay, has been a subject of study for decades, and will likely be so for many years to come. The optimistic prediction of Pereira et al. (1998a) that the Hg levels in the basin would return to background values in 20 years is clearly not a reality. Nearly 20 years since the publication, contamination is still observed, as is the gradient of contaminants from the effluent discharge site to other sites of the basin.

It was also observed that the halophyte *Halimione portulacoides* is able to accumulate several metals in its tissues, mainly in the BG tissues. Endophytic isolates from contaminated and non-contaminated sites revealed PGP activity in several in vitro assays, and *Pseudomonas* spp. isolates from contaminated sites showed ability to promote *in vivo* growth of the model plant *Arabidopsis thaliana*. This evidences the role that endophytic bacteria play in the endosphere of plants that intervene in accumulation of contaminants and in the phytoremediation of contaminated soils. This role has also been recently observed in other plants regarding metal contamination (e.g., Pan et al., 2017).

## Considerations on the description of novel bacterial species

Taxonomy and the description of new taxa have a reputation of being composed of strict rules. During the description of ten novel species from five different genera in the present work, it was observed that this statement is both true and untrue. On the one hand, classes of tests are required for the description of novel taxa, such as genomic sequence-based tests (e.g., 16S rRNA gene sequence), biochemical and phenotypical tests. In the case of specific taxa, e.g. those to be included in the family *Halomonadaceae*, an extensive list of tests exists for each class of tests. This is what makes the initial statement true: rules are in place so that researchers are required to perform a certain number and type of tests. On the other hand, the way these tests are performed is not as strict. A number of different methods may be used to assess, for example, the range and optima for temperature, pH and salt requirements: tests may be performed on liquid or solid media, the exact points of temperature/pH/salt to test are chosen by the researchers, as is the duration and several other aspects of the assays. To make matters worse, details of the conditions that the researchers use in their experiments are rarely described in the published papers. This makes comparison of such a simple test -- assessing range and optima of growth conditions -- a highly subjective matter. This subjectivity, however, is not common to all tests since, for some biochemical tests, researchers are required to assay their putative novel species, and closely related reference type strains in parallel, in the same conditions. Nevertheless, parallel testing is not a requirement for all biochemical and phenotypic tests. This drawback could begin to be surpassed if it was required of the researchers to describe all methodology in detail, since it is an essential part (if not *the* essential part) of strain comparison for taxa description purposes in the case of tests that are not required to be ran in parallel for novel and reference strains.

The genomic sequence-based criteria for strain comparison with the aim of species delineation have been chiefly based on the 16S rRNA gene sequence. However, since the sequence of this gene does not efficiently discriminate closely related species (de la Haba et al., 2010), its use for studying some taxa was deemed insufficient. In those cases, the information given by the 16S rRNA gene has been complemented with other genomic sequence-based approaches to aid in bacterial species delineation. Such approaches include the concatenation of sequences of a number of housekeeping genes that act as molecular clocks, as was performed for the description of novel taxa of *Microbacterium* and *Salinicola* in Chapter 4. Consequently, this MLSA approach uses more information, in terms of number of nucleotides, than the 16S rRNA gene sequence approach. As

bacterial genome sequencing becomes a reality for a larger number of research groups, the number of nucleotides that may be used for strain comparison is vastly greater. Genome comparison ought to be the next obvious step as input for strain comparison and to help delineate bacterial species. Accordingly, bioinformatics tools that allow for this comparison have been developed and optimized, to the point where even as little as 20 to 25 % of the genome of a strain is estimated to suffice in giving a stable result to be used as sequence-based criteria in species delineation (Richter & Rosselló-Móra, 2009).

The whole genome sequencing era should make it possible to have enough information that gives a highly reliable answer to, at least, which genus a strain belongs to. This determination could give way for other methods to be considered obsolete for these genus confirmation assays, such as listing polar lipids, fatty acids or determining dominant ubiquinones. Undeniably, these traits should be defined for a genus, for example at the moment of genus description. But perhaps their determination in the description of subsequent species could be considered redundant in determining something that is easily known with sufficient genome-based information. Additionally, common characteristics determined in *wet lab* context such as determination of G+C content and performing DNA-DNA hybridization ought to be replaced with *in silico* determination of these traits. These *in silico* determinations have relevant advantages when compared to their *wet lab* counterparts, namely (i) no need to use a specialized laboratory to perform the method, (ii) no need for sensitive and error prone methodologies, and (iii) the ability to build an additive database (Meier-Kolthoff et al., 2013).

The ever-changing sciences of taxonomy and systematics face new challenges with the advances in and greater availability of genome sequencing. These challenges should be met with comprehensive discussion so as to enhance the process of description of novel taxa and get as much information as possible from genome-based methods. It is of imperative importance to take these new techniques into account as soon as possible, since over 750 novel species of bacteria were published in the International Journal of Systematic and Evolutionary Microbiology in 2016 alone.

### **Final considerations on studies on endophytic bacteria**

As discussed in Chapters 3 and 5, culturable bacteria usually represent only a small fraction of the bacteria present in a community. In the present work, the efforts applied to isolate endophytic bacteria from *H. portulacoides* yielded a considerable number of isolates, when compared to similar types of studies in the literature. This may have been the result of using three different culture media to isolate bacteria, as it allowed coverage of nutritional needs of oligotrophs, of bacteria that require high salt contents, and of bacteria that grow in regular nutritive media. Amplicon-based culture-independent methods used to study communities present advantages in comparison to culture-dependent methods, especially when the depth of the information obtained is considered. Nevertheless, it is still highly relevant to study the culturable fraction. An assessment of the response of a community exposed to contaminants can be performed by analysing the culturable fraction from contaminated and non-contaminated samples, as discussed in Chapter 3. Additionally, having a collection of isolates allowed for a deeper characterization of a vast selection of traits, which is not yet achievable with the current tools in a culture-independent approach.

High-throughput sequencing allowed the detection of an additional dominant phylum that had not been observed in the culturable methodology. In this sense, the present work contributed to further show that culture-based methods do not suffice to have a comprehensive understanding of a bacterial community, even when multiple nutritional needs are considered in the culture media used.

There are obstacles in the analysis of endophytic bacterial communities using culture-independent methods. Primarily, the interference of host DNA in PCR amplification when using universal primers that target the bacterial 16S rRNA gene sequence. The majority of methods available to avoid this interference are based on processes that also induce loss of information of the bacterial fraction. The PNA blocker approach appears to be adequate in the sense that, in theory, there is no loss of information in the process of obtaining the community DNA sample, nor in the PCR amplification. In our case, using the PNA blockers allowed for an in-depth analysis of the diversity, in spite of a marked host DNA presence. This approach allowed us to carry out Illumina-based sequencing to accomplish our aim of performing a deep and extensive characterization of the bacterial diversity in the endosphere of *H. portulacoides*.

## Discussion of main and specific goals

The present work accomplished its main goal of characterizing the endophytic bacterial community of the salt marsh halophyte *H. portulacoides*.

Regarding the specific goals listed in Chapter 2, these were also achieved: in Chapter 3, a comprehensive collection of 665 endophytic isolates was obtained from three sites with different levels of metal(loid) contamination. This collection was subjected to PCR-fingerprinting, yielding 467 distinct strains, which were then explored for their taxonomy and relevant characteristics such as production of enzymes and presence of plant growth promotion (PGP) traits. Promising isolates for PGP were detected and further studied for their *in vivo* abilities in a model plant, specifically those belonging to the genus *Pseudomonas*. The endophytic bacteria collection was also explored for the presence of undescribed taxa and, in Chapter 4, ten novel bacterial species are described in detail according to taxa-specific regulations. Finally, the diversity of endophytic bacteria residing in healthy specimens of *H. portulacoides* was extensively explored by combining knowledge from culture-dependent (Chapter 3) and -independent methods (Chapter 5).

## Considerations regarding the proposed hypotheses

Four hypotheses were proposed in Chapter 2 and explored throughout the work.

The first hypothesis stated that the endosphere of the halophyte *H. portulacoides* is a hotspot for bacterial diversity, and this was in fact observed throughout the work. In Chapter 3, culture-based methodologies were applied to obtain a large collection of endophytic bacteria, and these efforts revealed that the culturable endosphere of *H. portulacoides* comprised four phyla, seven classes, forty families and seventy nine genera. Culture-independent methods used in Chapter 5 attempted at a deeper analysis of the diversity and allowed identification of 37 phyla, 85 classes, 182 families and 232 genera. Considering these numbers, the results supported the first hypothesis.

The second hypothesis stated that the presence of metal(loid)s influences the structure and diversity of the endophytic bacteria of *H. portulacoides*. This hypothesis was only tested for the culturable fraction of the endophytic bacteria of the halophyte, and it was indeed observed that the structure and diversity was significantly different in samples collected from metal(loid) contaminated and non-contaminated sites.

The third hypothesis stated that the endophytic bacterial community of the halophyte includes bacteria that have potential for promoting plant growth. This was observed in Chapter 3 where several isolates obtained in the culturable fraction presented one or more PGP traits, and where *Pseudomonas* spp. isolates also exhibited antagonistic properties against known pathogens, and the ability to induce *in vivo* plant elongation.

The fourth hypothesis stated that the endophytic bacterial community of the halophyte includes a plethora of novel bacteria taxa. Information based on partial sequences of the 16S rRNA gene of representative strains in the endophytic bacteria collection revealed 29 putative novel species, 7 of which represented putative novel genera. From these, ten novel species were described (Chapter 4) and are now (or will be soon) members of the known diversity in the domain Bacteria.

## Conclusions

The endosphere of the halophyte *Halimione portulacoides* is a hotspot for bacterial diversity which differs according to the presence of metal(loid)s in the sediment, and should be further explored for their potential to aid in phytoremediation. The analysis of this unexplored microbial hotspot also revealed a plethora of undescribed diversity, exposing a promising reservoir of new bacterial taxa.

## Future work

In the present work several of the obtained endophytic isolates exhibited an enormous potential for promoting plant growth. In Chapter 3, *Pseudomonas* spp. isolates were further analyzed for their PGP potential and *in vitro* experiments revealed the ability of the most promising isolates to promote growth of the model plant *Arabidopsis thaliana*. Future work should be performed in a similar manner with other promising isolates, so as to assess their ability to participate in tasks such as (i) promote growth of different plants, including those used in phytoremediation processes; (ii) increase tolerance to contaminants, especially for those isolates obtained from metal(loid) contaminated sites; (iii) increase tolerance to salt, since the isolates were collected from a salt-tolerant plant.

Several strains in the endophytic collection belonged to novel bacterial species. Ten novel species are described in the present work, however, many more remain obscured from the known bacterial diversity. Future work should be performed in order to unveil this diversity, and, during such an endeavor, the techniques used should include those which allow for cumulative knowledge to complement databases, such as genome sequencing, as discussed in Chapter 4 and in the present Chapter.

Finally, future work should include a culture-independent approach to analyze the bacterial endophytome from *H. portulacoides* sampled from sites with intermediate and high metal(loid) contamination. This would allow for a deeper analysis of such an endophytome and also allow for comparison with the results obtained in Chapter 5, potentially revealing a highly adapted community in a contaminated environment.

## References

- de la Haba, R. R., Arahal, D. R., Márquez, M. C. & Ventosa, A. (2010a).** Phylogenetic relationships within the family *Halomonadaceae* based on comparative 23S and 16S rRNA gene sequence analysis. *Int J Syst Evol Microbiol* **60**, 737-748.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P. & Göker, M. (2013).** Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**, 60.
- Pan, F., Meng, Q., Luo, S., Shen, J., Chen, B., Khan, K. Y., Japenga, J., Ma, X., Yang, X. & Feng, Y. (2017).** Enhanced Cd extraction of oilseed rape (*Brassica napus*) by plant growth-promoting bacteria isolated from Cd hyperaccumulator *Sedum alfredii* Hance. *Int J Phytorem* **19**, 281-289.
- Pereira, M. E., Duarte, A. C., Millward, G. E., Abreu, S. N. & Vale, C. (1998a).** An estimation of industrial mercury stored in sediments of a confined area of the Lagoon of Aveiro (Portugal). *Wat Sci Tech* **37**, 125-130.
- Richter, M. & Rosselló-Móra, R. (2009).** Shifting the genomic gold standard for the prokaryotic species definition. *PNAS* **106**, 19126-19131.